(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 December 2002 (19.12.2002)

PCT

(10) International Publication Number
WO 02/101086 A2

(51) International Patent Classification[7]: C12Q 1/68

(21) International Application Number: PCT/CA02/00848

(22) International Filing Date: 10 June 2002 (10.06.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/297,340 11 June 2001 (11.06.2001) US

(71) Applicant *(for all designated States except US)*: **1304854 ONTARIO LTD.** [CA/CA]; 3rd Floor, 1294 Balmoral Street, Thunder Bay, Ontario P7B 5Z5 (CA).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **BIRCH-MACHIN, Mark** [GB/GB]; 5 Emblehope Drive, Kingsmere, Gosforth, Newcastle-upon-Tyne NE3 4RW (GB). **DAKUBO, Gabriel, D.** [CA/CA]; 6273 Castille Court, Ottawa, Ontario K1C 1X4 (CA). **PARR, Ryan** [US/CA]; 1282 Hutton Park, Thunder Bay, Ontario P7G 1J4 (CA). **THAYER, Robert** [CA/CA]; RR 13, Site 9, Comp. 38, Thunder Bay, Ontario P7B 5E4 (CA). **NGOM, Alioune** [SN/CA]; 744 Sunset Avenue, Windsor, Ontario N9B 3B5 (CA). **TH'NG, John** [CA/CA]; Northwestern Ontario Regional Cancer Centre, 290 Munro Street, Thunder Bay, Ontario P7A 7T1 (CA).

(74) Agent: **TORYS LLP**; Suite 3000, 79 Wellington St. W., Box 270, TD Centre, Toronto, Ontario M5K 1N2 (CA).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/101086 A2

(54) Title: COMPLETE MITOCHONDRIAL GENOME SEQUENCES AS A DIAGNOSTIC TOOL FOR THE HEALTH SCIENCES

(57) Abstract: The examination of mutations in the entire mitochondrial genome is used as a diagnostic system for diseases such as prostate cancer, and non melanoma skin cancer. Characteristic mutations and rearrangements including, point mutations (transitions, transversions), deletions, inversions, duplications, recombinations, insertions or combinations thereof in the mitochondrial genome are used as early indicators of prostate cancer, and non melanoma skin cancer. Moreover, the 4977bp, or "common deletion" as well as other associated mutations and/or deletions are used as a measure of aging.

Complete Mitochondrial Genome Sequences as a Diagnostic Tool for the Health Sciences

Technical Field of the Invention

This invention is related to the field of mitochondrial genomics. In particular it is related to mutations in the mitochondrial genome and their utility as an indicator of the very genesis of disease.

Background of the Invention

The current mega-trend in the biological sciences is the human genome project, and commercial exploitation of the data. However, there is an exceptional limitation to the use and implementation of this information as the data is not specific at the level of the individual. Incredibly the data is from only a few individuals, hardly representative of the variation present in human populations, rendering the data useful in general applications only. The staggering complexity of the human genome makes application on an individual basis impractical. To sequence completely one human nuclear genome the U.S. Department of Energy and the National Institute of Health have invested 2.5 billion dollars since 1988 (http://www.ornl.gov/hgmis/project/budget.html).

Mitochondrial Genome

The mitochondrial genome is a compact yet critical sequence of nucleic acid. The mitochondrial genome codes for enzyme subunits necessary for cellular respiration. Mitochondrial DNA, or "mtDNA", is a minuscule genome of nucleic acid at 16,569 base pairs (bp Anderson et al., 1981; Andrews et al., 1999) in contrast to the immense nuclear genome of 3.3 billion bp. Its genetic complement is astronomically smaller than that of its nuclear cell mate (0.0005%); however, communication or chemical signalling, routinely occur (Sherratt et al., 1997). Moreover, specific nuclear components are responsible for maintenance and integrity of mitochondrial sequence (Croteau et al., 1999). When these nuclear areas are rendered non-functional by nuclear rearrangements indicative of potential disease, then mutations begin to appear in mtDNA sequences. In addition, specific mitochondria may be identified for intracellular destruction by deletions prompted by somatic mutations in the mitochondrial genome. This theoretical mechanism may serve as an indication of impending disease as well. About 3,000 genes are required to make a

mitochondrion, with only thirty-seven of these coded by the mitochondrial genome, indicating heavy mitochondrial dependence on nuclear loci (Naviaux, 1997).

The essential role of mtDNA is the generation of the cellular fuel, adenosine triphosphate (ATP), which fires cellular metabolism. Significantly, the mitochondrial genome is dependent on seventy nuclear encoded proteins to accomplish the oxidation and reduction reactions necessary to this vital function, in addition to the thirteen polypeptides supplied by the mitochondrial genome (Leonard and Shapira, 1997). Different tissues and organs depend on oxidative phosphorylation to a varied extent. Moreover, mutations in the mitochondrial genome are associated with a variety of chronic, degenerative diseases (Gattermann et al. 1995). Diseases related to defective oxidative phosphorylation (OXPHOS) appear to be closely linked to mtDNA mutations (Byrne, 1992). Consequently as OXPHOS diminishes due to increased severity of mtDNA mutations, organ specific energetic thresholds are exceeded which give rise to a variety of clinical phenotypes.

Recently, Fliss et al. (2000) found, in primary tumors from lung and bladder cancer, a high frequency of mtDNA mutations which were predominantly homoplasmic in nature, indicating that the mutant mtDNA was dominant in the malignant cells. Point mutations and deletions would appear to be the non-programmed but unavoidable side effect of oxygen free radical damage to the membrane and genome of mitochondria (Miquel et al. 1992). This theory is plausible because not only is the mitochondrial genome lacking protective histones, but also is vulnerable to oxidative damage being found near the oxygen generating inner mitochondrial membrane. Moreover, as mtDNA has a compact genome and lacks introns, deleterious events are thus likely to affect a coding sequence resulting in a biochemical dysfunction. This dysfunction will further increase cellular oxidative stress which will lead to nuclear as well as mtDNA damage, thereby increasing the potential for a cell to enter into the cancer process (Penta et al., 2001). In this respect, research indicates that with increasing age there is an increase in mtDNA damage (Cortopassi & Wang 1995) and a subsequent decline in respiratory function (Miquel et al. 1992) leading to eventual cell death.

MtDNA as a Diagnostic Tool

MtDNA sequence dynamics are important diagnostic tools. Mutations in mtDNA are often preliminary indicators of developing disease, often associated with nuclear mutations, and act as biomarkers specifically related to disease, such as but not limited to: tissue damage and cancer from smoking and exposure to second hand tobacco smoke (Lee et al., 1998; Wei, 1998); longevity, based on accumulation of mitochondrial genome mutations beginning around 20 years of age and increasing thereafter (von Wurmb, 1998); metastatic disease caused by mutation or exposure to carcinogens, mutagens, ultraviolet radiation (Birch-Machin, 2000); osteoarthritis; cardiovascular, Alzheimer, Parkinson disease (Shoffner et al., 1993; Sherratt et al., 1997;Zhang et al, 1998); age associated hearing loss (Seidman et al., 1997); optic nerve degeneration and cardiac dysrhythmia (Brown et al., 1997; Wallace et al., 1988); chronic progressive external exophthalmoplegia (Taniike et al., 1992); atherosclerosis (Bogliolo et al., 1999); papillary thyroid carcinomas and thyroid tumours (Yeh et al., 2000); as well as others (e.g. Naviaux, 1997; Chinnery and Turnbull, 1999;).

Specifically, these alterations include point mutations (transitions, transversions), deletions (one base to thousands of bases), inversions, duplications, (one base to thousands of bases), recombinations and insertions (one base to thousands of bases). In addition, specific base pair alterations, deletions, or combinations of are associated with early onset of prostate, skin, and lung cancer, as well as aging (e.g. Polyak et al., 1998), premature aging, exposure to carcinogens (Lee et al., 1998), etc.

Since mtDNA is passed to offspring exclusively through the ovum, it is imperative to understand mitochondrial sequences through this means of inheritance. The sequence of mtDNA varies widely between maternal lineages (Ward et al., 1991), hence mutations associated with disease must be clearly understood in comparison to this variation. For example, a specific T to C transition noted in the sequence of several individuals, associated with a specific cancer, could in reality be natural variation in a maternal lineage widespread in a given particular geographical area or associated with ethnicity. For example, Native North Americans express an unusually high frequency of adult onset diabetes. In addition, all North American Natives are genetically characterized by five

basic maternal lineages designated A, B, C, D, and X (Schurr et al., 1990; Stone and Stoneking, 1993; Smith et al., 1999). Lineage A is distinguished by a simple point mutation resulting in a *Hae III* site at bp 663 in the mitochondrial genome, yet there is no causative relationship between this mutation and the adult onset of diabetes. In addition, even within lineage clusters there is sequence variation.

Outside of the specific markers associated with a particular lineage there is more intrapopulation variation than interpopulation sequence variation (Easton et al., 1996; Ward et al., 1991, 1993;) This divergence must be understood for optimal identification of disease associated mutations, hence a maternal line study approach (Parsons et al., 1997), mimicking the strengths of a longitudinal design (i.e. subject tracking over a substantial period of time), must be used to identify mutations directly associated with disease, as opposed to mutations without disease association. Moreover, particular substances, such as second hand tobacco smoke, low levels of asbestos, lead, all known mutagens and at low levels in many environments, may be the cause of specific point mutations, but not necessarily a disease specific marker. Hence, a substantial mtDNA sequence database is a clear prerequisite to accurate forecasting of potential disease as a natural process, or through exposure to causative agents. Furthermore, the entire molecule must be sequenced for its full information content. The entire suite of point mutations (transitions, transversions), deletions (one base to thousands of bases), inversions, duplications, (one base to thousands of bases), recombinations and insertions (one base to thousands of bases) must be characterized as a whole over the entire mitochondrial genome. This ensures that all possible information available in the mitochondrial genome is captured. Although the genome of cytoplasmic mitochondria (16,569bp) has been sequenced at an individual level, like its nuclear counterpart, the mitochondrial genome has not been sequenced at a population level for use as a diagnostic tool.

Recently mitochondria have been implicated in the carcinogenic process because of their role in apoptosis and other aspects of tumour biology (Green & Reed, 1998, Penta et al., 2001), in particular somatic mutations of mtDNA (mtDNA) have been observed in a number of human tumours (Habano et al. 1998; Polyak et al. 1998; Tamura et al. 1999; Fliss, et al. 2000). These latter findings were made more interesting by the claims that the

particular mtDNA mutations appeared to be homoplasmic (Habano et al. 1998; Polyak et al.1998; Fliss, et al. 2000). Additionally researchers have found that ultraviolet radiation (UV) is important in the development and pathogenesis of non-melanoma skin cancer (NMSC) (Weinstock 1998; Rees, 1998) and UV induces mtDNA damage in human skin (Birch-Machin, 2000a).

Moreover, through time, mitochondrial sequence loses integrity. For example, the 4977bp deletion increases in frequency with age (Fahn et al., 1996). Beginning at age 20, this deletion begins to occur in small numbers of mitochondria. By age 80, a substantial number of molecules have been deleted. This deletion characterizes the normal aging process, and as such serves as a biomarker for this process. Quantification of this aging process may allow medical or other interventions to slow the process.

This application of mitochondrial genomics to medicine has been overlooked because mtDNA has been used primarily as a tool in population genetics and more recently in forensics; however, it is becoming increasingly evident that the information content of mtDNA has substantial application in the field of medical diagnostics. Moreover, sequencing the entire complement of mtDNA was a laborious task before the recent advent of high capacity, high-throughput robotic DNA sequencing systems. In addition, population geneticists were able to gather significant data from two highly variable areas in the control region; however, these small regions represent a small portion of the overall genome, less than 10%, meaning that 90% of the discriminating power of the data is left unused! Significantly, many disease associated alterations are outside of the control region. The character of the entire genome should be considered to include all sequence information for accurate and highly discriminating diagnostics.

Non-Melanoma Skin Cancer

Human non-melanoma skin cancer (NMSC) is the commonest cancer in many Caucasian populations (Weinstock, 1998; Rees, 1998). The majority of these tumours are basal cell carcinoma (BCC) and squamous cell carcinoma (SCC). BCCs are locally invasive and can cause significant morbidity but rarely metastasis. SCCs show significant metastatic potential and the occurrence of multiple NMSCs in patients with

immunosuppression causes significant management problems (Rees, 1998). While there are no clinically identified pre-malignant lesions for BCC, some SCCs are thought to arise from precursor lesions, namely actinic keratoses (AKs) or areas of Bowen's disease (in situ carcinoma)(Rees, 1998).

SCCs show loss of heterozygosity affecting several chromosomes which suggests the involvement of several tumour suppressor genes in their development. Interestingly, in AKs, an equal or greater degree of genetic loss is observed in these precursor lesions compared to SCCs (Rehman et al. 1994; Rehman et al. 1996). This is important for the proposed invention because it suggests that other mechanisms, in addition to inactivation of tumour suppressor genes, are likely to be involved in the development of SCCs.

A role for mitochondria in tumourigenesis was originally hypothesised when tumour cells were found to have an impaired respiratory system and high glycolytic activity (Shay & Werbin, 1987). Recent findings elucidating the role of mitochondria in apoptosis (Green & Reed, 1998) together with the high incidence of homoplasmic mtDNA mutations in colon cancer (Habano et al. 1998; Polyak et al. 1998, reviewed in Penta et al., 2001), primary tumours of the bladder, neck and lung (Fliss et al. 2000), and gastric tumours (Tamura et al. 1999), further support this hypothesis. Furthermore, it has been proposed that these mitochondrial mutations may affect the levels of reactive oxygen species (ROS) which have been shown to be highly mitogenic (Polyak et al. 1998; Li et al. 1997).

Previous studies by the inventors and others have shown that mutations in mtDNA and the associated mitochondrial dysfunction is an important contributor to human degenerative diseases (Birch-Machin et al. 1993; Chinnery et al. 1999; Birch-Machin et al. 2000b). This is because the mitochondrial genome is particularly susceptible to mutations due to the high amounts of ROS produced in this organelle coupled with the lack of protective histones and a low rate of mtDNA repair (Pascucci et al. 1997; Sawyer & van Houten; LeDoux et al.1999) compared to the nucleus. lndeed, the mutation rate for mtDNA is around ten times higher than that of nuclear DNA (Wallace,1994). Most of the mtDNA mutations identified in the recent human tumour studies have indicated possible exposure to ROS derived mutagens. This is important for the investigation of mtDNA mutations in

NMSC because there is recent evidence for the direct involvement of UV induced ROS in the generation of mtDNA deletions in human skin cells (Berneburg et al. 1999, Lowes et al., 2002). In addition, the major determinant of NMSC in individuals without protective pigmentation or genetic predisposition is UV (Weinstock, 1998). The putative precursor lesions of SCCs are also found predominantly on constant sun-exposed sites. This is important because work by the Birch-Machin laboratory has shown distinct differences between the incidence of mitochondrial DNA damage in skin taken from different sun-exposed body sites. The vast majority of the damage is found on constant sun-exposed sites (Krishanan et al., 2002).

One of the inventors was the first to quantitatively show that UV exposure induces mtDNA damage (Birch- Machin et al. 1998). MtDNA as a molecular marker was used to study the relation between chronological aging and photo aging in human skin. A 3-primer quantitative PCR method was used to study the changes in the ratio of the 4977 bp-deleted to wild type mtDNA in relation to sun exposure and chronological age of human skin. There was a significant increase in the incidence of high levels (i.e. >1 %) of the 4977bp-deleted mtDNA in sun-exposed (27%, [27/100]) compared with sun-protected sites (1.1% [1/90]) (Fishers exact test, $P<0.0001$). Deletions or mutations of mtDNA may therefore be useful as a marker of cumulative ultraviolet radiation exposure.

Furthermore, a study using a South-Western Blot approach involving monoclonal antibodies against thymine dimers, provided direct evidence for the presence of UV-induced damage in purified mtDNA (Ray et al. 1998).

Quantification of a single deletion alone, however, may not provide a reliable UV bio-marker because it represents one of many possible deletions or combinations and other associated mutations (Birch-Machin, 2000). Recent work from the inventors' research group has used a long extension PCR (LX-PCR) technique to amplify the entire mitochondrial genome in order to determine the whole deletion spectrum of mtDNA secondary to UV exposure (Ray et al. 2000). Long PCR analysis of 71 split skin samples, where the epidermis is separated from the underlying dermis, was performed in relation to sun exposure. There was a significant increase in the number of deletions with increasing UV exposure in the epidermis (Kruskal-Wallis test, $p=0.0015$). The findings in the

epidermis are not confounded by any age-dependent increases in mtDNA deletions also detected by the long PCR technique. The large spectrum of identified deletions highlights the ubiquitous nature and the high mutational load of mtDNA associated with UV exposure. Compared to the detection of single deletions using competitive PCR, the study shows that long PCR is a sensitive technique and may therefore provide a more comprehensive, although not quantitative, index of overall mtDNA damage in skin. The studies by one of the inventors described above clearly show that mtDNA is a significant target of UV and this together with the role of mitochondrial in skin disease has been recently reviewed (Birch-Machin, 2000).

The pigmentation of human hair and skin which is the major co-variant of UV sensitivity and human skin cancer has been investigated. These investigations have centred on the association of variants of the melanocortin 1 -receptor gene and sun-sensitivity of individuals and populations (Smith et al. 1998; Healy et al. 1999; Flanagan et al. 2000; Healy et al. 2000; Harding et al. 2000; Flanagan et al., 2002) relating to skin cancer susceptibility. However, these studies have not addressed population-level variation in mtDNA sequences in association with particular skin types and/or hair colour.

One of the questions which remains largely unanswered by the recent studies of mtDNA mutations in human tumours is the incidence of deletions of the mitochondrial genome in relationship to these tumours. This is an important question to answer because a preliminary study of a single patient in human skin has shown differences in the incidence of the common mtDNA deletion between several tumours (AKs and SCCs) and normal skin (Pang et al. 1994). As well, the inventors' own preliminary data shows an increased number of mtDNA deletions in tumours compared to normal skin. Finally, Birch-Machin and others have shown that the incidence of mtDNA deletions, as well as duplications, increases with increasing UV exposure (Berneburg et al. 1999; Birch-Machin et al. 1998;Ray et al. 1998; Ray et al. 1999; Ray et al. 2000), Lindsey et al., 2001; Birch-Machin et al., 2001; Lowes et al., 2002, Krishnan et al., 2002).

Apart from the questions relating to tumour progression other vital questions remain largely unanswered by the recent studies of mtDNA in human tumours (Habano et al. 1998; Fiiss et al. 2000). Firstly, due to technical limitations, it is not clear whether the

mtDNA mutations are truly homoplasmic, as varying levels of heteroplasmy may indicate important disease transitions as well (Habano et al. 1998; Polyak et al. 1998; Fliss, et al. 2000); secondly, apart from one study (Tamura et al. 1999) the incidence of mtDNA deletions and their role as potential biomarkers for NMSC was not investigated. Researchers have looked at the common deletion and ignored the rest of the 100 or so deletions. As well, investigators have been focused on identification of mutations, rather than their quantification. It is important to assess accurately in a quantitative manner the incidence of deletions because of the threshold effect of mtDNA damage on ATP production and consequently cell function. In addition, deletions are difficult to characterize. Long PCR is typically used which produces a ladder of deletions which then have to be characterized.

Current diagnosis of NMSC is pathological evaluation of excised tissue. Accordingly, there is a need for an early marker of UV-induced DNA damage which predisposes an individual to NMSC. There is also a need for a genetic-based diagnostic tool which allows for early detection and is diagnostically accurate.

Prostate Cancer

Prostate cancer is a frequently diagnosed solid tumour that most likely originates in the prostate epithelium (Huang et al. 1999). In 1997, nearly 10 million American men were screened for prostate specific antigen (PSA), the presence of which suggests prostate cancer (Woodwell, 1999). Indeed, this indicates an even higher number of men screened by an initial digital rectal exam (DRE). In the same year, 31 million men had a DRE (Woodwell, 1999). Moreover, the annual number of newly diagnosed cases of prostate cancer in the United States is estimated at 179,000 (Landis et al., 1999). It is the second most commonly diagnosed cancer and second leading cause of cancer mortality in Canadian men. In 1997 prostate cancer accounted for 19,800 of newly diagnosed cancers in Canadian men (28%) (National Cancer Institute of Canada). It is estimated that 30% to 40% of all men over the age of forty-nine (49) have some cancerous prostate cells, yet only 20% to 25% of these men have a clinically significant form of prostate cancer (SpringNet – CE Connection, internet, www.springnet.com/ce/j803a.htm). Prostate cancer exhibits a wide variety of histological behaviour involving both erogenous and exogenous factors, i.e.

socio-economic situations, diet, geography, hormonal imbalance, family history and genetic constitution (Konishi et al. 1997; Hayward et al. 1998).

From a risk standpoint *familial* and *hereditary* prostate cancers are not considered synonymous terms. Familial cancers refer to the incidences within a family, but are not inherited. This form accounts for up to 25% of prostate cancers (Walsh & Partin, 1997). Hereditary refers to a subtype of prostate cancer with a Mendelian inheritance of a predisposing gene(s) and accounts for approximately 9% of reported cases. A positive family history of prostate cancer for this disease suggests that these predisposing gene(s) play an important role in prostate cancer development and progression. Recently, susceptibility genes on chromosomes 1 and X have been identified as predisposing men to prostate cancer, providing greater insight into the etiology of hereditary cancer (Berthon et a]. 1998; Xu et al. 1998).

Prostate cancer prognosis mainly depends on the tumour stage and grade at diagnosis. Only localized prostate cancer can be cured by radical treatment. Standard detection still relies on digital rectal examination, PSA testing and histopathologic examination of prostatic biopsied tissues. Biopsy of a mass is used to confirm malignancy, it is not an early detection technique. Unfortunately, some early tumours are impossible to identify during rectal exams. PSA tests have a specificity of 60 to 70% and a sensitivity of 70 to 80% (personal communication, Dr. Sunil Gulavita, Northwestern Ontario Cancer Centre). A newer technique which refines diagnosis for tumours of common histologic grade is ploidy-DNA analysis employing flow cytometry (Shankey et al. 1995); however, this technique measures chromosomal changes that are only apparent in later stages of cancer development and is not sufficiently sensitive for the detection of minor alterations in DNA structure or chromosomal inversions, or reciprocal trans-locations in early cancers. The invention focuses on early detection since prognosis is heavily dependent on the stage of disease at diagnosis.

Our understanding of genetic abnormalities in prostate cancers is scanty. Research into prostate cancer has focussed on the development of knowledge in the following areas: 1) proto-oncogenes (Buttyan et al. 1987); 2) tumour suppressor genes (p53, p73, KAI1 and

MMACI/PTEN; Dong et al. 1995; Cairns et al. 1997) and 3) telomere/telomerase activity in metastasis. Up-regulation of telomerase and amplification of telomeric DNA in prostate cells may provide effective markers for diagnosis. Moreover, telomeres may serve as a site for therapy (Ozen et al. 1998). A number of groups have provided evidence for a "prostate cancer gene" in the short arm of chromosome 1 (Berthon et al. 1998). More work is needed to identify the specific locus within this region. It has been suggested that this marker is only one of several possible genes predisposing men to familial prostate cancer. Other studies have shown possible marker loci on the X chromosome (Xu et al. 1998). If some prostate cancers are polygenic, then mtDNA becomes an important diagnostic tool since it may be difficult to identify and understand the interplay between all associated nuclear genes in such cases.

Certainly, a key issue in prostate cancer research is to identify molecular markers that can effectively determine and distinguish tumour progression. Molecular markers may be able to discriminate between those cases of prostate neoplasmy which will proceed rapidly to metastatic disease and those with little chance of resulting in tumour development. Up to the present research has focused primarily on the secrets hidden within the nuclear genome; however, the much smaller mtDNA genome seems to act as a barometer for events in the nucleus and as such provides a means for the early detection of human prostate cancer (Zeviani et al. 1990). Importantly, in this respect, mitochondria have been implicated in the carcinogenic process because of their role in apoptosis and other aspects of tumour biology (Green & Reed 1998). In particular, somatic mutations of mtDNA have been observed in a number of human tumours (Polyak et al. 1998, Tamura et al. 1999, Fliss et al. 2000). However, previous studies have been exclusively cross-sectional as they have not considered the clonal nature of mtDNA in maternal lines. These limited cross-sectional studies merely show the mutation at one time point. This may or may not give an accurate link between a mutation and the corresponding disease state. Cross-sectional studies employing a maternal line have the advantage of tracking a mutation in mtDNA over time and thus mimic the strength of a longitudinal design. Mutations which are common population variants, as opposed to mutations associated with disease, can both be identified.

Aging

Aging consists of an accumulation of changes with time both at the molecular and cellular levels; however, the specific molecular mechanisms underlying the aging process remain to be elucidated. In an attempt to explain the aging process, mitochondrial genomes in older subjects are compared to the genomes of younger subjects from the same maternal lineage. One deletion associated with aging is known as the common deletion, or 4977-bp deletion. Aging research has been limited to this common deletion and polymorphisms in the control region. For a clear understanding of these mutations, the entire genome must be analyzed. Other deletions are seen in Table 1 adapted from Wei, 1992.

Table 1

Deletions size (bp)	References
4977	Cortopassi and Arnheim, 1990; Ikebe et al., 1990; Linnane et al., 1990; Corral-Debrinski et al., 1991; Yen et al., 1991; Torii et al., 1992; Zhang et al., 1992
7436	Corral-Debrinski et al., 1991; Hattori et al., 1991 Hsieh and Wei, 1992
3610	Katayama et al., 1991
6063	Hsieh and Wei, 1992 Yen et al., 1992
5827	Zhang et al., 1992
6335	Zhang et al., 1992
7635	Zhang et al., 1992
7737	Zhang et al., 1992
7856	Zhang et al., 1992
8041	Zhang et al., 1992
8044	Zhang et al., 1992
5756	Zhang et al., 1992

Oxygen free radicals, a normal by product of ATP production, are a probable cause of this deletion, which increases in frequency with age. Existing literature demonstrates a strong association between mtDNA (mtDNA) mutations, chronological age, and the overall aging process in postmitotic tissues such as muscle and brain; however, comparative maternal line studies are needed to discriminate between aging associated mutational events and those mutations without an aging association.

In recent years a variety of chronic degenerative diseases have been shown to result from mutations in mtDNA (Gatterman et al. 1995). Diseases related to defective OXPHOS appear to be closely linked to mtDNA mutations (Byrne, 1992). Furthermore, it has been shown that these myopathies are often associated with the common deletion of 4977-bp of

the mitochondrial genome (Liu et al. 1997). This large deletion has also been found, at heteroplasmic levels, in various tissues of normal aging persons and is consistent with the Mitochondrial Theory of Aging (Harman, 1981). This is manifest through an increase in the deletion frequency (Cortopassi & Wang, 1995) and a subsequent decline in respiratory function (Miquel et al. 1992) resulting in eventual cell death in old age. The early detection of a predisposition to a disease or disorder presents the best opportunity for medical intervention, as early genetic diagnosis may improve the prognosis for a patient.

Previous studies employing a cross-sectional design have established an association or cause and effect relationship between mtDNA mutations, deletions, and/or combinations of such and aging; however, in order to obtain accurate data the age specific deletion and/or mutation rate must be determined concisely. Attributing mutations to the aging process as opposed to a particular disease at the population level is vital. This information is imperative to an understanding of how mtDNA damage accrues over time. Moreover, the consequences of these particular mutations, their frequencies, and associations in the temporal aspects of aging must be known in order to forecast and eventually slow aging at the molecular level. Researchers have not yet determined this rate, which requires evaluation of population data through maternal lines. Accordingly, there is a need for a biomarker which tracks the aging process.

Accordingly, there is a need for a simple, straightforward system of monitoring the vast nuclear genome for mutations which indicate early stage cancer, aging or other human diseases with a DNA component. There is also a need for a simple diagnostic system for non-melanoma skin cancer, prostate cancer, lung cancer and aging linked to defects in the mitochondrial genome. There is a need for a diagnostic system which differentiates between mutations in mtDNA which cause disease, and those which simply represent variation within and between populations.

Summary of the Invention

An object of the present invention is to provide a simple, straightforward system for monitoring the vast nuclear genome for early transitions associated with cancer, aging, and other human diseases with a DNA component.

In an embodiment of the present invention a small biological sample which includes tissue or fluid samples such as urine, prostate fluid, skin cells, or saliva is taken from an individual. These samples are examined, using any suitable method including histological examination, to identify cells demonstrating disease morphology. Using any suitable method, including without limitation; laser capture, identified cells demonstrating disease morphology are recovered from the sample and the mtDNA therefrom is sequenced, followed by comparison to a database of known mitochondrial sequences associated with both health and disease.

In a preferred embodiment, the entire mitochondrial genome is sequenced at a population level to determine the variation of mtDNA sequences associated with disease.

In an additional embodiment, the presence of mutation progression may signal the beginning and continuing development of disease. Mutation load may also indicate progression or disease state.

In a preferred embodiment, mtDNA sequences from prostate massage fluid are compared to a mtDNA sequence database of normal, transitory, and metastatic mtDNA sequences clearly associated with prostate cancer. This comparative data set is based on studies of maternal lines, and other normal maternal line variation present in the population stored in a maternal line database affording a lucid picture of mtDNA mutations clearly associated with disease, as opposed to variation present in mitochondrial lineages existing in the general population.

There may be specific maternal lineages which indicate a predisposition to disease.

In another embodiment, mtDNA sequences from suspected non-melanoma skin cancers are compared to a mtDNA sequence database of normal and mtDNA sequences clearly associated with non-melanoma skin cancer.

According to an aspect of the present invention, there is provided a method of detecting in a subject containing mtDNA the genesis or progression of disease comprising obtaining a biological sample from the subject, extracting DNA from the biological sample, and detecting the presence of mutations in the mtDNA. The step of detecting the presence of mutations is selected from the group consisting of sequencing the mtDNA, amplifying the mtDNA by PCR, South-Western blotting, denaturing HPLC, hybridization to microarrays, gene chips or biochips, molecular marker analysis or combinations thereof. Further, the mtDNA of the biological sample is compared to a database, the database containing data of mutations associated with the mtDNA sequences of non-disease and disease associated mitochondrial genomes.

According to an aspect of the present invention, there is provided a method of detecting in a human subject the presence of a disease comprising obtaining a biological sample from the human subject, extracting DNA from the biological sample, detecting mutations in the mitochondrial DNA of the biological sample, and comparing the mitochondrial DNA sequence of the biological sample to a database, the database containing data of mutations associated with the mitochondrial DNA sequences of non-disease and disease associated mitochondrial genomes Mutation rates of mitochondria DNA associated with a specific disease may be an important indicator of disease development and prognosis. This may allow specific identification of disease stage, improving disease definition resulting in better disease intervention and specific therapy application.

In yet another embodiment, increasing the sensitivity for heteroplasmy detection increases the early identification capacity of the test.

In addition, the invention may be used to monitor the progression of disease by watching important sites targeted by metastasis.

According to another aspect of the present invention, there is provided a method of determining a predisposition to a disease or disorder indicated by mutations in a mitochondrial DNA sequence comprising: obtaining a biological sample from the human

subject, extracting DNA from the biological sample, detecting mutations in the mitochondrial DNA of the biological sample, and comparing the mitochondrial DNA sequence of the biological sample to a database, the database containing data of mutations associated with the mitochondrial DNA sequences of individuals who are predisposed to the disease or disorder, and individuals who are not predisposed to the disease or disorder.

In a preferred embodiment, a DNA microarray is used in determining the sequence of the mitochondrial DNA. Other technologies can also be used. For example, direct sequencing of a subset, or the complete human genome, SNaP shot™, SNP detection, real time PCR or other methods as is standard in the art.

According to a further aspect of the present invention, there is provided a method for assessing the status of the aging process of a human subject comprising obtaining a biological sample from the human subject, extracting DNA from the biological sample, detecting mutations in the mitochondrial DNA of the biological sample, and comparing the mitochondrial DNA sequence of the biological sample to a database, the database containing data of mutations of TDNA associated with aging.

The step of detecting the presence of mutations in the mtDNA can be selected from: sequencing the mtDNA, amplifying mtDNA by PCR, Southern, Northern, Western, South-Western blot hybridizations, denaturing HPLC, hybridization to microarrays, biochips or gene chips, molecular marker analysis or a combination of any of the above.

According to yet another aspect of the present invention, there is provided a database containing a plurality of human mitochondrial DNA sequences, the mitochondrial DNA sequences selected from the group of normal control sequences associated with non-disease states, sequences associated with the presence of disease or sequences indicative of the predisposition to disease.

According to yet another aspect of the present invention, there is provided a kit for diagnosis of a disease comprising a disposable chip, microarray, means for holding the disposable chip, means for extraction of mitochondrial DNA and means for access to a database of mitochondrial DNA sequences.

According to yet another aspect of the present invention there is provided a method of diagnosing a disease in a patient comprising hybridizing a nucleic acid sample obtained from mitochondrial DNA to an array comprising a solid substrate and a plurality of nucleic acid members, wherein each member is indicative of the presence of a disease, wherein each nucleic acid member has a unique position and is stably associated with the solid substrate, and wherein hybridization of said nucleic acid sample to one or more nucleic acid members comprising said array is indicative of the presence of prostate cancer.

According to yet another aspect of the present invention there is provided a kit for determining predisposition to a disease comprising a disposable chip, microarray, means for holding the disposable chip, means for extraction of DNA and means for access to a database of mitochondrial DNA sequences.

According to another aspect of the present invention, there is provided a method of determining a predisposition to or developing symptoms of a disease or disorder indicated by mutations in a mitochondrial DNA sequence comprising obtaining a biological sample from the human subject, extracting mitochondrial DNA from the biological sample, sequencing the mitochondrial DNA of the biological sample, and comparing the mitochondrial DNA sequence of the biological sample to a database, the database containing population-level data of mutations associated with the mtDNA sequences of non-disease and disease associated mitochondrial genomes.

According to yet another aspect of the present invention, there is provided a method of diagnosing non-melanoma skin cancer in a patient comprising: hybridizing a nucleic acid sample obtained from mitochondrial DNA to an array comprising a solid substrate and a plurality of nucleic acid members, wherein each member is indicative of non-melanoma cancer, wherein each nucleic acid member has a unique position and is stably associated with the solid substrate, and wherein hybridization of said nucleic acid sample to one or more nucleic acid members comprising said array is indicative of the presence of non-melanoma skin cancer. Alternatively, non-specific mutations may reach a threshold effect beyond which cancer develops. In a similar manner, prostate cancer can also be diagnosed.

According to another aspect of the present invention, there is provided a method of detecting heteroplasmy in a subject containing mtDNA comprising obtaining a biological sample from the subject; extracting DNA from the biological sample; and performing denaturing HPLC on the sample.

According to another aspect of the present invention, there is provided a method of detecting mutations associated with disease in a subject containing mtDNA comprising: obtaining a biological sample from the subject, extracting DNA from the biological sample, detecting the presence of mutations in the mtDNA, and comparing the mtDNA of the biological sample to a database, the database containing data of common population variants in non-disease and disease associated mitochondrial genomes.

Detailed Description of the Invention

The method of the present invention can be used to diagnose diseases linked to mtDNA. The method of the present invention provides for amplification of the mitochondrial genome of an individual from a biological sample, sequencing a portion of the mitochondrial genome, preferably the entire mitochondrial genome of the individual using any known means. Denaturing high performance liquid chromatography (DHPLC) may also be used to rapidly screen many samples. DHPLC can focus on hotspots of mutations. DHPLC is more sensitive than automated sequencing in terms of detecting mutations, 2% heteroplasmy, compared with 20-25% for ordinary sequencing. Methods for detecting lower levels of heteroplasmy (<2%) may also be developed.

As used herein, the "presence" of a mutation in mtDNA includes heteroplasmic mutations and, therefore, it is contemplated that there may be additionally the presence of some normal mtDNA in a sample in which the mutated DNA is present.

As used herein, "actinic kerotoses" means proposed precursor epidermal lesion of a squamous cell carcinoma.

As used herein, "aging" refers to an accumulation of changes with time, both at the molecular and cellular levels.

As used herein, "alleles" means one of several alternative forms of a given DNA sequence occupying a specific place on a chromosome.

As used herein, "attaching" or "spotting" refers to a process of depositing a nucleic acid onto a solid substrate to form a nucleic acid array such that the nucleic acid is irreversibly bound to the solid substrate via covalent bonds, hydrogen bonds or ionic interactions.

As used herein, "basal cell carcinoma" means a type of cancer of skin cells.

As used herein, "Bowen's disease" means in situ epidermal carcinoma.

As used herein, "diagnostic" or "diagnosing" means using the presence or absence of a mutation or combination of mutations as a factor in disease diagnosis or management. The detection of the mutation(s) can be a step in the disease state diagnosis.

As used herein, "disease" includes a disorder or other abnormal physical state.

As used herein, "disease associated mitochondiral genomes" means genomes containing mutations indicative or otherwise associated with a particular disease.

As used herein, "database" means an electronic storage system (computer based using standard industry software) which will have the capacity to store and provide retrievable information that will enable researchers to rapidly determine the structure of the nucleotide sequences. The database will also store descriptive information about those individuals who provide the biological samples. This descriptive information will include health status and other pertinent indices which may be correlated to the biological sample.

As used herein, "deletions" means removal of a region of DNA from a contiguous sequence of nucleic acids, where once a deletion has occurred, the gap is repaired by rejoining of the ends. Deletions can range in size from one base to thousands of bases or larger.

As used herein, "duplications" means when a specific sequence of DNA is copied and inserted behind or forward of the original copy one or more times or elsewhere in the genome.

As used herein, "heteroplasmy" is defined by the ratio of mutant: to wild type mtDNA molecules, where 100% mutant mtDNA is termed "homoplasmic". Heteroplasmic mutations are those mutations which occur in some, but not all of the copies of the mitochondrial genome.

As used herein, "homoplasmy" means all mitochondrial sequences are identical.

As used herein, "inversions" refers to when a length of DNA is excised and reinserted in reverse orientation.

As used herein, "maternal inheritance" means mitochondria which are inherited through the cytoplasm of the ovum.

As used herein, "maternal line" refers to the clonal sequence of mitochondrial DNA as passed down through successive generations from the mother.

As used herein, "mitochondria" means a eukaryotic cytoplasmic organelle that generates ATP for cellular processes.

As used herein, "mutation" encompasses any change in a DNA sequence from the wild type sequence, including without limitation point mutations, transitions, insertions, transversions, translocations, deletions, inversions, duplications, recombinations or combinations thereof.

As used herein, "mutation load" refers to an increase in mutations in mtDNA which eventually leads to compromised function of the involved gene or the entire genome.

As defined herein, a "nucleic acid array" refers to a plurality of unique nucleic acids attached to one surface of a solid support at a density exceeding 20 different nucleic acids/cm^2 wherein each of the nucleic acids is attached to the surface of the solid support in a non-identical preselected region. In one embodiment, the nucleic acid attached to the surface of the solid support is DNA. In a preferred embodiment, the nucleic acid attached to the surface of the solid support is cDNA. In another preferred embodiment, the nucleic acid attached to the surface of the solid support is cDNA synthesized by polymerase chain reaction (PCR). Preferably, a nucleic acid array according to the invention, comprises nucleic acids of at least 150 nucleotides in length. Preferably, a nucleic acid array comprises nucleic acids of less than 6,000 nucleotides in length. More preferably, a nucleic acid array comprises nucleic acids of less than 500 nucleotides in length. In one embodiment, the array comprises at least 500 different nucleic acids attached to one surface of the solid support. In another embodiment, the array comprises at least 10 different nucleic acids attached to one surface of the solid support. In yet another embodiment, the array comprises at least 10,000 different nucleic acids attached to one surface of the solid support. The term "nucleic acid", as used herein, is interchangeable with the term "polynucleotide".

As used herein, a " nucleic acid target" or "a target nucleic acid" is defined as a nucleic acid capable of binding to a nucleic acid member of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a nucleic acid target may include natural (i. e., A, G, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, the bases in nucleic acid probe may be joined by a linkage other than a phosphodiester bond, so long as it does not interfere with hybridization. Thus, nucleic acid targets may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. Preferably, the nucleic acid targets are derived

from human tissue or fluid extracts. More preferably, the nucleic acid targets are single- or double-stranded DNA, RNA, or DNA-RNA hybrids synthesized from human tissue of fluid extracts.

As used herein, "nucleus" means the most conspicuous organelle in the eucaryotic cell, contains all of the chromasomal DNA.

As used herein, "PSA Test" means prostate-specific antigen test; an antigen found in blood that may be indicative of cancer of the prostate.

As used herein, "point mutation" means the change of a single nucleotide in DNA.

As used herein, "polymorphism" means sequence variation in a population of alleles or mtDNA genomes.

As used herein, "precursor lesions" means a DNA mutation, or combinations thereof, indicating potential disease association.

As used herein, "predisposed to a disease" or a "predisposition to a disease" means that individuals are at higher risk for developing the disease or disorder or are at higher risk for early onset of the disease or disorder than the average individual, due to the presence or absence of mutations which are associated with the disease or disorder.

As used herein, "preselected region", "predefined region", or "unique position" refers to a localized area on a substrate which is, was, or is intended to be used for the deposit of a nucleic acid and is otherwise referred to herein in the alternative as a "selected region" or simply a "region." The preselected region may have any convenient shape, e.g., circular, rectangular, elliptical, wedge-shaped, etc. In some embodiments, a preselected region is smaller than about 1 cm^2, more preferably less than 1 mm^2, still more preferably less than 0.5 mm^2, and in some embodiments about 0.125 to 0.5 mm^2.

As used herein, "somatic mutation" means a change in DNA sequence after fertilization.

As used herein, "solid substrate" or "solid support" refers to a material having a rigid or semi-rigid surface. The terms "substrate" and "support" are used interchangeable herein with the terms "solid substrate" and "solid support". The solid support may be biological, non-biological, organic, inorganic, or a combination of any of these, existing as particles, strands, precipitates, gels, sheets, tubing, spheres, containers, capillaries, pads, slices, films, plates, slides, etc. Often, the substrate is a silicon or glass surface, (poly)tetrafluoroethylene, (poly)vinylidendifluoride, polystyrene, polycarbonate, a charged membrane, such as nylon 66 or nitrocellulose, or combinations thereof. In a preferred embodiment, the solid support is glass. Preferably, at least one surface of the substrate will be substantially flat. Preferably, the surface of the solid support will contain reactive groups, including, but not limited to, carboxyl, amino, hydroxyl, thiol, or the like. In one embodiment, the surface is optically transparent.

As used herein, "squamous cell carcinoma" means a type of cancer of skin cells.

As used herein, "stably associated" refers to a nucleic acid that is irreversibly bound to a solid substrate to form an array via covalent bonds, hydrogen bonds or ionic interactions such that the nucleic acid retains its unique preselected position relative to all other nucleic acids that are stably associated with an array, or to all other preselected regions on the solid substrate under conditions wherein an array is analyzed (i.e., hybridization and scanning).

A "statistically significant" number of mitochondrial DNA sequences is determined by or through the use of standard chi-square statistical algorithms using or determining observed versus expected scores.

As used herein, "transitions" means substitution of like nitrogenous bases, pyrimidine to pyrimidine, purine to purine. A mutation in which one pyrimidine is substituted by the other, or in which one purine is substituted by the other.

As used herein, "transversions" means substitution of unlike nitrogenous bases, purine to pyrimidine, pyrimidine to purine. A mutation in which a purine is substituted or replaced by a pyrimidine or vice versa.

MtDNA and diagnosis of specific diseases

In an embodiment of the present invention, methods are provided for monitoring aging and diagnosing specific diseases such as prostate cancer and non-melanoma skin cancer through comparisons of mtDNA sequences. Diagnosing diseases such as prostate cancer with mtDNA, rather than nuclear DNA has several advantages. Firstly, mtDNA, a less complex genome, is easily understood at an individual and population level, hence a large mtDNA database with normal and disease associated genomes renders individual diagnosis extremely accurate. Accordingly, variation, in relationship to disease, is understood. Secondly, mtDNA has a 10-fold higher mutation rate than nuclear DNA (Wallace 1992). Nuclear rearrangements, suggestive of preliminary disease, are rapidly communicated to mitochondria, where they appear as somatic mutations. Thirdly, mtDNA has a maternal inheritance pattern, and is essentially clonal in that all mitochondria begin with the same mtDNA sequence, hence variation from this clonal condition is easily detected. Additionally, mtDNA does not show convincing evidence of recombination, thus any alterations in sequence are a somatic event. Any one mitochondrion harboring a mutation(s) is in a sense 'recessive' as a consequence of there being many mitochondrial genomes (2-10 copies) per mitochondrion, and many mitochondria per cell (500-2,000). Moreover, mitochondrial genomes can tolerate very high levels (up to 90%) of mitochondria with damaged genomes. This happens through complementation by the remaining wild type mtDNA (Chomyn et al. 1992). However, mutated genomes have a replicative advantage over wild type genomes because they are usually smaller (Hayashi et al. 1991), hence there is clonal expansion of mutated mtDNA (Brierley et al. 1998), suggesting that unlike nuclear genes, there is little or no selection against cells harboring mtDNA mutations. Because of this elevated mutation rate, mutations and/or deletions that appear in mtDNA are maintained through the life span of the cell and may serve as a record of exposures to various mutagens. The integrity of mtDNA is maintained by nuclear repair mechanisms, and a defect at these loci has been suggested to result in an autosomal

dominant disorder associated with multiple mitochondrial deletions (Zeviani et al. 1990). Consequently, mtDNA may function as an early warning sentinel of early nuclear events related to a variety of cancers or other diseases. Finally, the mitochondrial genome can be sequenced and monitored for mutations on an individual basis.

Diagnosis of Non-Melanoma Skin Cancer

In a preferred embodiment of the invention, a system for early diagnosis of mtDNA changes in non-melanoma skin cancer (NMSC) and their precursor lesions indicative of solid tumour development is provided. The particular changes, such as the common deletion and associated mutations, and the incidence of as yet uncharacterised deletions in mtDNA serve as reliable bio-markers of potential skin cancer. The mutation fingerprint of the entire mtDNA genome in human NMSC and its precursor lesions is determined. Thus mtDNA changes are established as an early bio-marker of human skin cancer and its precursor lesions. Denaturing HPLC can then be used to assess low levels of heteroplasmy at the sequences of interest. This approach can also provide an insight into the development of early changes in other human tumours.

Diagnosis of prostate cancer

In another embodiment of the invention, a system for diagnosis of prostate cancer is provided. Age related accumulation of mtDNA defects might predispose an individual to the appearance of certain clinical disorders such as prostate cancer which is prevalent in middle age and older men. In a preferred embodiment, routine prostate cancer screening takes place through mitochondrial genome sequencing from prostate massage fluid. The presence of epithelial cells transformed into cancer cells, can be determined through amplification of mtDNA from prostate massage fluid, eclipsing current diagnostic techniques such as digital rectal examination and PSA. Recently Fliss et al. (2000) identified mutated mtDNA in urine samples of patients with bladder cancer. Similar findings in prostate massage fluid provide a non-invasive early detection method for prostate cancer. Different types of prostate cancer can be diagnosed, as well as differentiating between aggressive, fast growing cells in younger patients in contrast to prostate cancer as a whole.

Assessment of mutations associated with aging

The system and method of the present invention may be used to assess aging, based on the increasing frequency of mutations such as the "common deletion" of 4977-bp and other mutations of the mitochondrial genome (Liu et al. 1997). This information, in conjunction with health survey data, allows crucial statistical discrimination between separate causes resulting in the same mutation/deletion. Fortunately mtDNA is inherited exclusively through the ovum and is essentially clonal in nature (Van De Graaff & Fox, 1995). This permits carefully controlled studies of mutations/deletions within maternal lines through several generations to determine a reliable age related deletion frequency. This information may be used to develop treatment methods which slow the aging process.

Collection of samples

Biological samples can be collected by any known means, whether for the purpose of constructing a mtDNA sequence database, or performing a diagnostic test on an individual. Samples destined for database generation include, but are not limited to: tumour banks, maternal lineage studies involving affected and unaffected individuals from the same maternal lineage, as well as maternal lineage studies from groups or populations with high frequencies of specific disease such as, but not limited to: skin and prostate cancer, assessment of health status and aging. For example, FTA® Gene Cards® may be used to collect and archive biological samples. Suitable samples include any tissue or body fluid derived from mesothelium, epithelium, or endothelium. Such tissues and fluids include, but are not limited to blood, sputum, buccal cells, saliva, prostate massage fluid, sweat, bone, hair, lymph tissue, cervical smears, breast aspirate, fecal matter, ejaculate, menstrual flow and biopsy tissue. Preferably, approximately 100 µl of blood, 100 µg to 25 mg of solid tissue is sampled. In the case of suspected skin cancer, skin cells or tissue, (from normal, NMSC and precursor lesions) is taken from skin biopsy or a routine suction blistering technique. Where a disease is suspected, primary care physicians, oncologists or other practitioners, may extract both normal and suspected disease tissue from the patient.

For samples of tumours such as prostate or skin, replicate cross-sections (5 microns) of micro-dissected paraffin embedded tissues are de-paraffinized prior to one slide being stained with hematoxylin and eosin (HE), with the replicate stained with methyl

green (MG), as is standard in the art. HE stains are graded by a pathologist for normal, precursor, and applicable grades of tumour progression. Replicate MG slides are used for laser capture, according to manufacturers recommendations (Arcturus) of graded cells.

Extraction of mtDNA

Extraction of DNA may take place using any method known in the art, followed by sequencing of the mitochondrial genome, as described in Current Protocols in Molecular Biology.

Sequencing of MtDNA

PCR

Polynucleotide sequences of the invention can be amplified by the polymerase chain reaction (PCR). PCR methods are well-known to those skilled in the art. PCR requires the presence of a nucleic acid to be amplified, two single stranded oligonucleotide primers flanking the sequence to be amplified, a DNA polymerase, deoxyribonucleoside triphosphates, a buffer and salts. The method of PCR is well known in the art. PCR is performed as described in Mullis and Faloona, 1987, Methods Enzymol., 155: 335, herein incorporated by reference.

In general, PCR is performed using template DNA (at least 1fg; more usefully, 1-1000 ng) and at least 25 pmol of oligonucleotide primers. A typical reaction mixture includes: 2μl of DNA, 25 pmol of oligonucleotide primer, 2.5 μl of 10X PCR buffer 1 (Perkin-Elmer, Foster City, CA), 0.4 μl of 1.25 μM dNTP, 0.15 μl (or 2.5 units) of Taq DNA polymerase (Perkin Elmer, Foster City, CA) and deionized water to a total volume of 25 μl. Mineral oil is overlaid and the PCR is performed using a programmable thermal cycler.

The length and temperature of each step of a PCR cycle, as well as the number of cycles, are adjusted according to the stringency requirements in effect. Annealing temperature and timing are determined both by the efficiency with which a primer is expected to anneal to a template and the degree of mismatch that is to be tolerated. The ability to optimize the stringency of primer annealing conditions is well within the

knowledge of one of moderate skill in the art. An annealing temperature of between 30°C and 72°C is used. In general, initial denaturation of the template molecules normally occurs at between 92°C and 99°C for 4 minutes, followed by 20-40 cycles consisting of denaturation (94-99°C for 15 seconds to 1 minute), annealing (temperature determined as discussed above; 1-2 minutes), and extension (72°C for 1 minute). The final extension step is generally carried out for 4 minutes at 72°C, and may be followed by an indefinite (0-24 hour) step at 4°C.

DNA Sequencing

Any known means to sequence the mitochondrial genome may be used. Preferably, mtDNA is amplified by PCR prior to sequencing. PCR products can be sequenced directly or cloned into a vector which is then placed into a bacterial host. Examples of DNA sequencing methods are found in Brumley, R. L. Jr. and Smith, L.M., 1991, Rapid DNA sequencing by horizontal ultrathin gel electrophoresis, Nucleic Acids Res. 19:4121-4126 and Luckey, J.A., et al, 1993, High speed DNA sequencing by capillary gel electrophoresis, Methods Enzymol. 218: 154-172. The combined use of PCR and sequencing of mtDNA is described in Hopgood, R., et al, 1992, Strategies for automated sequencing of human mtDNA directly from PCR products, Biotechniques 13:82-92 and Tanaka, M. et al, 1996, Automated sequencing of mtDNA, Methods Enzymol. 264: 407-421

Deletion Analysis and Detection

A preferable approach is the long extension PCR (LX-PCR) technique using the Expand Long Template PCR system (Boehringer Mannheim). Using the LX-PCR technique, which has been established and validated in the Birch-Machin laboratory (Ray et al. 2000), there is the opportunity to rapidly screen for the whole spectrum of mtDNA deletions as opposed to the incidence of a single deletion.

A semi-quantitative PCR method (Corral-Debrinski *et al* 1991) can be used to estimate the proportion of the $mtDNA^{4977}$ deletion in the total mtDNA.

In addition, Southern Blot and probing technology labeled with isotopes or any other technique as is standard in the art may be used for deletion detection as well.

Sequencing of PCR products

Any known means may be used to sequence the PCR products. Preferably, the entire DNA sequence is characterized by di-deoxy sequencing using ABI Big Dye Terminator™ technology and a series of 72 overlapping primers each for heavy and light strands. Sequencing occurs on one, several, or a combination of ABI platforms such as the 310, 3100, or 3700. Sequencing reactions are performed according to manufacturer's recommendation.

Mutational analysis of the mitochondrial genome using denaturing high performance liquid chromatography (DHPLC)

Prior to sequencing of the mitochrondrial genome and identification of mutational hotspots, DHPLC can be used to rapidly screen mutations in many samples. This technique provides greater sensitivity in identification of low levels of heteroplasmy. It cannot detect homoplasmic changes but will complement traditional sequencing. Apart from the homoplasmic mutations recently identified in tumours, the vast majority of reported mtDNA mutations are heteroplasmic (Chinnery et al. 1999). These heteroplasmic mtDNA changes result in the formation of heteroduplexes after PCR amplification of the mtDNA. Rapid screening for heteroplasmic mtDNA mutations is determined using the relatively new technique of denaturing high performance liquid chromatography (DHPLC) (Oefner & Underhill, 1998). This technique has recently been used to rapidly screen and identify whole mtDNA genomes for heteroplasmic point mutations down to levels <5% (Van den Bosch et al. 2000).

The DHPLC may be performed on the WAVE™ DNA Fragment Analysis System (Transgenomic, Omaha, USA) which provides a fully automated screening procedure. The same technology can be used to screen for mtDNA heteroplasmic mutations. Preferably, the entire mtDNA genome is amplified by PCR in 13 overlapping fragments using two different PCR conditions as described by van den Bosch et al. (2000). The 1-2 kb PCR products are digested into fragments of 90-600bp and resolved at their optimal melting temperature. Mutations are represented as two peaks and mutations with low percentages, such as <2% heteroplasmy as a 'shoulder' in the peak.

DNA sequencing can also take place using a microarray, as is known in the art (Chee et al. 1996).

Data Analysis

Once sequenced, normal and disease associated mtDNA sequences are archived for comparison in a database. Resequencing devices, micro-array technology, integrated microfluidic amplification and analysis systems, high-speed, high-throughput, mutation detection, and other methods may all be used with the methods of the present invention.

Data obtained from the sequencing of the individual mitochondrial genome is compared to population level data. The data is obtained through obtaining samples and sequencing mtDNA as described above. Preferably, the database contains information from maternal line studies. The population level data is maintained in a database. Any suitable database can be used.

Preferably, a multidimensional evaluation research database of clinical and biological data is used, which provides the bio-informatics infrastructure necessary for the collection, processing and dissemination of information amassed by the laboratories involved in this venture. The database is a centralized electronic system which links networks resulting in a dynamic and powerful resource.

The database may be accessed through any known means, and preferably through a secure Internet pathway. Preferably, the database is developed using an e-commerce algorithm, built on a server and deployed using an application server which supports a high volume of concurrent users through optimized performance and scalability features. A separate "web" server can provide the foundation of the web-site architecture since it can serve as the central point through which all content, applications, and transactions must flow before reaching users.

Data mining algorithms known in the art are used to discover patterns, clusters and models from data (SAS 2000). Moreover, intelligent algorithms and methods will be

developed for: occurrence of mutation and mutation rates, patterns of mutations for disease detection, information retrieval, and other complex sequence analysis software.

Nucleic Acid Members and Probes

The invention provides for nucleic acid members and probes that bind specifically to a target nucleic acid sequence. The target nucleic acid sequence is a nucleic acid or a region of a nucleic acid that is to be detected, as indicative of disease such as prostate cancer, non-melanoma skin cancer and the like. The target nucleic acid sequences to be analyzed using a microarray of the invention are preferably derived from human tissue or fluid samples. The invention provides for target nucleic acid sequences comprising RNA or nucleic acid corresponding to RNA, (i.e., cDNA), or DNA. Nucleic acid members are stably associated with a solid support to comprise an array according to the invention. The nucleic acid members may be single or double stranded, and may be a PCR fragment amplified from cDNA.

The invention also provides for polynucleotide sequences comprising a probe. As used herein, the term "probe" refers to an oligonucleotide which forms a duplex structure with a sequence in the target nucleic acid, due to complementarity of at least one sequence in the probe with a sequence in the target region. The probe may be labeled, according to methods known in the art. A probe according to the invention may be single or double stranded.

Diagnostic devices

The invention includes diagnostic devices such as biochips, gene chips or microarrays used to diagnose specific diseases or identify specific mutations. All sequenced mitochondrial genomes are assessed to create a consenus structure of the base pair arrangement and are assigned a prohibiting index for proportion of base pair deletions and mutations associated with a particular disease or disorder. The diagnostic arrangement is then used to create biochips, gene chips, or microarrays.

Once sequences associated with particular diseases, disease states or disorders are identified, hybridization of mtDNA to an array of oligonucleotides can be used to identify particular mutations. Any known method of hybridization may be used. Preferably, an array is used, which has oligonucleotide probes matching the wild type or mutated region, and a control probe. Commercially available arrays such as microarrays or gene chips are suitable. These arrays contain thousands of matched and control pairs of probes on a slide or microchip, and are capable of sequencing the entire genome very quickly. Review articles describing the use of microarrays in genome and DNA sequence analysis is available at www.gene-chips.com.

Microarray

Polynucleotide arrays provide a high throughput technique that can assay a large number of polynucleotides in a sample comprising one or more target nucleic acid sequences. The arrays of the invention are useful for gene expression analysis, diagnosis of disease and prognosis of disease (e.g., monitoring a patient's response to therapy, drug screening, and the like).

Any combination of the polynucleotide sequences of mtDNA indicative of disease, aging, or other health related mutations are used for the construction of a microarray.

The target nucleic acid samples to be analyzed using a microarray are derived from any human tissue or fluid which contains adequate amounts of mtDNA, as previously described, preferably prostate massage fluid, solid tumours, blood, or urine. The target nucleic acid samples are contacted with polynucleotide members under hybridization conditions sufficient to produce a hybridization pattern of complementary nucleic acid members/target complexes.

Construction of a microarray

The microarray comprises a plurality of unique polynucleotides attached to one surface of a solid support, wherein each of the polynucleotides is attached to the surface of the solid support in a non-identical preselected region. Each associated sample on the array comprises a polynucleotide composition, of known identity, usually of known sequence, as

described in greater detail below. Any conceivable substrate may be employed in the invention.

The array is constructed using any known means. The nucleic acid members may be produced using established techniques such as polymerase chain reaction (PCR) and reverse transcription (RT). These methods are similar to those currently known in the art (see e.g. PCR Strategies, Michael A. Innis (Editor), et al. (1995) and PCR: Introduction to Biotechniques Series, C. R. Newton, A. Graham (1997)). Amplified polynucleotides are purified by methods well known in the art (e.g., column purification). A polynucleotide is considered pure when it has been isolated so as to be substantially free of primers and incomplete products produced during the synthesis of the desired polynucleotide. Preferably, a purified polynucleotide will also be substantially free of contaminants which may hinder or otherwise mask the binding activity of the molecule.

In the arrays of the invention, the polynucleotide compositions are stably associated with the surface of a solid support, wherein the support may be a flexible or rigid solid support.

Any solid support to which a nucleic acid member may be attached may be used in the invention. Examples of suitable solid support materials include, but are not limited to, silicates such as glass and silica gel, cellulose and nitrocellulose papers, nylon, polystyrene, polymethacrylate, latex, rubber, and fluorocarbon resins such as TEFLON™.

The solid support material may be used in a wide variety of shapes including, but not limited to slides and beads. Slides provide several functional advantages and thus are a preferred form of solid support. Due to their flat surface, probe and hybridization reagents are minimized using glass slides. Slides also enable the targeted application of reagents, are easy to keep at a constant temperature, are easy to wash and facilitate the direct visualization of RNA and/or DNA immobilized on the solid support. Removal of RNA and/or DNA immobilized on the solid support is also facilitated using slides.

The particular material selected as the solid support is not essential to the invention, as long as it provides the described function. Normally, those who make or use the invention will select the best commercially available material based upon the economics of cost and availability, the expected application requirements of the final product, and the demands of the overall manufacturing process.

Numerous methods are used for attachment of the nucleic acid members of the invention to the substrate (a process referred as spotting). For example, polynucleotides are attached using the techniques of, for example U.S. Pat. No. 5,807,522, which is incorporated herein by reference for teaching methods of polymer attachment. Alternatively, spotting is carried out using contact printing technology.

The amount of polynucleotide present in each composition will be sufficient to provide for adequate hybridization and detection of target polynucleotide sequences during the assay in which the array is employed. Generally, the amount of each nucleic acid member stably associated with the solid support of the array is at least about 0.1 ng, preferably at least about 0.5 ng and more preferably at least about 1 ng, where the amount may be as high as 1000 ng or higher, but will usually not exceed about 20 ng. Where the nucleic acid member is "spotted" onto the solid support in a spot comprising an overall circular dimension, the diameter of the "spot" will generally range from about 10 to 5,000 μm, usually from about 20 to 2,000 μm and more usually from about 50 to 1000 μm.

Control polynucleotides may be spotted on the array and used as target expression control polynucleotides and mismatch control nucleotides to monitor non-specific binding or cross-hybridization to a polynucleotide in the sample other than the target to which the probe is directed. Mismatch probes thus indicate whether a hybridization is specific or not. For example, if the target is present the perfectly matched probes should be consistently brighter than the mismatched probes. In addition, if all central mismatches are present, the mismatch probes are used to detect a mutation.

Target preparation

The targets for the microarrays, are derived from human fluid or tissue samples. It may be desirable to amplify the target nucleic acid sample prior to hybridization. One of

skill in the art will appreciate that whatever amplification method is used, if a quantitative result is desired, care must be taken to use a method that maintains or controls for the relative frequencies of the amplified polynucleotides. Methods of "quantitative" amplification are well known to those of skill in the art. For example, quantitative PCR involves simultaneously co-amplifying a known quantity of a control sequence using the same primers. This provides an internal standard that may be used to calibrate the PCR reaction. The high density array may then include probes specific to the internal standard for quantification of the amplified polynucleotide. Detailed protocols for quantitative PCR are provided in PCR Protocols, A Guide to Methods and Applications, Innis et al., Academic Press, Inc. N.Y., (1990). Other suitable amplification methods include, but are not limited to polymerase chain reaction (PCR) (Innis, et al., PCR Protocols. A guide to Methods and Application. Academic Press, Inc. San Diego, (1990)), ligase chain reaction (LCR) (see Wu and Wallace, Genomics, 4: 560 (1989), Landegren, et al., Science, 241: 1077 (1988) and Barringer, et al., Gene, 89: 117 (1990), transcription amplification (Kwoh, et al., Proc. Natl. Acad. Sci. USA, 86: 1173 (1989)), and self-sustained sequence replication (Guatelli, et al., Proc. Nat. Acad. Sci. USA, 87: 1874 (1990)).

The invention provides for labeled target or labeled probe. Any analytically detectable marker that is attached to or incorporated into a molecule may be used in the invention. An analytically detectable marker refers to any molecule, moiety or atom which is analytically detected and quantified. Detectable labels suitable for use in the present invention include any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels in the present invention include biotin for staining with labeled streptavidin conjugate, magnetic beads (e.g., DynabeadsTM), fluorescent dyes (e.g., fluorescein, texas red, rhodamine, green fluorescent protein, and the like), radiolabels (e.g., ^{3}H, ^{125}I, 35S, ^{14}C, or ^{32}P), enzymes (e.g., horseradish peroxidase, alkaline phosphatase and others commonly used in an ELISA), and colorimetric labels such as colloidal gold or colored glass or plastic (e.g., polystyrene, polypropylene, latex, etc.) beads. Patents teaching the use of such labels include U.S. Pat. Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241.

Means of detecting such labels are well known to those of skill in the art. Thus, for example, radiolabels may be detected using photographic film or scintillation counters, fluorescent markers may be detected using a photodetector to detect emitted light. Enzymatic labels are typically detected by providing the enzyme with a substrate and detecting the reaction product produced by the action of the enzyme on the substrate, and colorimetric labels are detected by simply visualizing the colored label.

The labels may be incorporated by any of a number of means well known to those of skill in the art. However, in a preferred embodiment, the label is simultaneously incorporated during the amplification step in the preparation of the sample polynucleotides. Thus, for example, polymerase chain reaction (PCR) with labeled primers or labeled nucleotides will provide a labeled amplification product. In a preferred embodiment, transcription amplification, as described above, using a labeled nucleotide (e.g. fluorescein-labeled UTP and/or CTP) incorporates a label into the transcribed polynucleotides. Alternatively, a label may be added directly to the original polynucleotide sample (e.g., mRNA, polyA mRNA, cDNA, etc.) or to the amplification product after the amplification is completed. Means of attaching labels to polynucleotides are well known to those of skill in the art and include, for example nick translation or end-labeling (e.g. with a labeled RNA) by kinasing of the polynucleotide and subsequent attachment (ligation) of a polynucleotide linker joining the sample polynucleotide to a label (e.g., a fluorophore).

In a preferred embodiment, the target will include one or more control molecules which hybridize to control probes on the microarray to normalize signals generated from the microarray. Labeled normalization targets are polynucleotide sequences that are perfectly complementary to control oligonucleotides that are spotted onto the microarray as described above. The signals obtained from the normalization controls after hybridization provide a control for variations in hybridization conditions, label intensity, "reading" efficiency and other factors that may cause the signal of a perfect hybridization to vary between arrays.

Hybridization conditions

Polynucleotide hybridization involves providing a denatured probe or target nucleic acid member and target polynucleotide under conditions where the probe or target nucleic acid member and its complementary target can form stable hybrid duplexes through complementary base pairing. The polynucleotides that do not form hybrid duplexes are then washed away leaving the hybridized polynucleotides to be detected, typically through detection of an attached detectable label. It is generally recognized that polynucleotides are denatured by increasing the temperature or decreasing the salt concentration of the buffer containing the polynucleotides. Under low stringency conditions (e.g., low temperature and/or high salt) hybrid duplexes (e.g., DNA:DNA, RNA:RNA, RNA:DNA, cDNA:RNA and cDNA:DNA) will form even where the annealed sequences are not perfectly complementary. Thus specificity of hybridization is reduced at lower stringency. Conversely, at higher stringency (e.g., higher temperature or lower salt) successful hybridization requires fewer mismatches. Methods of optimizing hybridization conditions are well known to those of skill in the art (see, e.g., Laboratory Techniques in Biochemistry and Molecular Biology, Vol. 24: Hybridization With Polynucleotide Probes, P. Tijssen, ed. Elsevier, N.Y., (1993)).

Following hybridization, non-hybridized labeled or unlabeled polynucleotide is removed from the support surface, conveniently by washing, thereby generating a pattern of hybridized target polynucleotide on the substrate surface. A variety of wash solutions are known to those of skill in the art and may be used. The resultant hybridization patterns of labeled, hybridized oligonucleotides and/or polynucleotides may be visualized or detected in a variety of ways, with the particular manner of detection being chosen based on the particular label of the test polynucleotide, where representative detection means include scintillation counting, autoradiography, fluorescence measurement, calorimetric measurement, light emission measurement and the like.

Image Acquisition and Data Analysis

Following hybridization and any washing step(s) and/or subsequent treatments, as described above, the resultant hybridization pattern is detected. In detecting or visualizing the hybridization pattern, the intensity or signal value of the label will be not only be detected but quantified, by which is meant that the signal from each spot of the hybridization will be measured and compared to a unit value corresponding to the signal

emitted by a known number of end labeled target polynucleotides to obtain a count or absolute value of the copy number of each end-labeled target that is hybridized to a particular spot on the array in the hybridization pattern.

Methods for analyzing the data collected from hybridization to arrays are well known in the art. For example, where detection of hybridization involves a fluorescent label, data analysis can include the steps of determining fluorescent intensity as a function of substrate position from the data collected, removing outliers, i.e., data deviating from a predetermined statistical distribution, and calculating the relative binding affinity of the test polynucleotides from the remaining data. The resulting data is displayed as an image with the intensity in each region varying according to the binding affinity between associated oligonucleotides and/or polynucleotides and the test polynucleotides.

Following detection or visualization, the hybridization pattern is used to determine quantitative information about the genetic profile of the labeled target polynucleotide sample that was contacted with the array to generate the hybridization pattern, as well as the physiological source from which the labeled target polynucleotide sample was derived. By genetic profile is meant information regarding the types of polynucleotides present in the sample, e.g. in terms of the types of genes to which they are complementary, as well as the copy number of each particular polynucleotide in the sample.

Diagnostic or Prognostic Tests

The invention provides for diagnostic tests for detecting diseases. The invention also provides for prognostic tests for monitoring a patient's response to therapy. According to the method of the invention, the presence of disease or the patient's response to therapy is detected by obtaining a fluid or tissue sample from a patient. A sample comprising nucleic acid is prepared from the fluid or tissue sample. The nucleic acid extracted from the sample is hybridized to an array comprising a solid substrate and a plurality of nucleic acid members, wherein each member is indicative of the presence of disease or a predisposition to a disease or disorder. According to this diagnostic test, hybridization of the sample comprising nucleic acid to one or more nucleic acid members on the array is indicative of disease, a predisposition to a disease or disorder, or in the case of a prognostic test, indicative of a patient's response to therapy.

Other utilities for the present invention, such as that described above and in the following examples, will be readily apparent to those skilled in the art.

The present invention is more particularly described in the following examples which are intended as illustrative only since numerous modifications and variations will be apparent to those skilled in the art.

Example 1: Prostate Tumours

Following acquisition of prostate fluid or surgery to remove prostate tumours, biopsy slides are prepared to identify transforming or cancerous cells. Laser Capture Microdissection (LCM) microscopy is used to isolate cells that are either normal, benign, or malignant from the tissue section. Procurement of diseased cells of interest, such as precancerous cells or invading groups of cancer cells is possible from among the surrounding heterogeneous cells.

Total DNA extraction from each of these cells was purified according to a modification of the protocol outlined by Arcturus Engineering Inc. DNA was extracted from cells with a 50µl volume of 1 mg/ml proteinase K (PK), in lOmM Tris pH 8.0, O.1mM EDTA pH 8.0, and 0.1% Tween 20, at 42°C overnight. Following incubation overnight at 42°C the tubes were removed from the incubation oven. The samples were microcentrifuged for 5 min at 640O rpm(2000 x g). The CapSure™ was removed from the tube and discarded. The tube was incubated at 95°C for 10 minutes (PK is inactivated) and then cooled to room temperature. 5-50µl of the sample was used for PCR amplification.

Following purification, individual samples are amplified, by LX-PCR using the appropriate primers for hypervariable region 1 (HV1), hypervariable region 2 (HV2) and the entire 12S region. These PCR products are then sequenced using high throughput methods as is well known in the art.

Example 2: Duplications in the non-coding region of mtDNA from sun-exposed skin

DNA was extracted from tissue samples as described in Example 1, with the use of DNeasy™ kit supplied by Qiagen. A "back to back" primer methodology was used to investigate the incidence of tandem duplications in the non-coding region (NCR) in relation to sun-exposure. 32 age-matched, split human skin samples, from sun-exposed (n=24) and sun-protected body sites (n=10) were investigated.

The following duplication primers from Brockington *et al* 1993 and Lee *et al* 1994 were used:

C	L336	AAC ACA TCT CTG CCA AAC CC	20 mer
D	H335	TAA GTG CTG TGG CCA GAA GC	20 mer
E	L467	CCC ATA CTA CTA ATC TCA TC	20 mer
F	H466	AGT GGG AGG GGA AAA TAA TG	20 mer

Primers pairs C/D and E/F are 'back to back' at the site of two separate sets of direct repeats in the non-coding region. As a result they only generate a product if a duplication is present at these points. Products generated are 260 bp and/or less common 200bp variant. Modified PCR conditions are: 100ng total cellular DNA, 200μM dNTPs, 2.5 U HotStarTaq polymerase and PCR buffer (Qiagen, Uk), 25 pmoles of primers: one cycle of 94°C for 4 minutes, 36 cycles of 94 °C x 1 minute, 55°C x 1 minute, 72°C x 1 minute and one cycle of 72°C x 7 minutes.

An increased incidence of duplications with increasing sun-exposure was observed, with duplications identified in 10/24 but 0/10 samples from sun-exposed and sun-protected skin respectively (Fisher's exact test, p=0.01 5) (Birch-Machin and Krishnan 2001). The sizes of the most frequent duplications were 200 and 260 base pairs. Interestingly these same samples also contained high levels (>1 %) of the 4977bp common mtDNA deletion as determined by an established quantitative 3-primer PCR assay described in Example 6.

Example 3: Mutation fingerprint of mtDNA in human NMSC and its precursor lesions

DNA was extracted from human skin tissue samples as described in Example 1, with the use of DNeasy™ by Qiagen Using specific primers, mtDNA is amplified by PCR

and following DNA sample preparation (Qiagen), mutations are identified by automated sequencing (PE Applied Biosystems) using BigDye™ Terminator Cycle sequencing. This methodology is described in Healy et al. 2000; Harding et al. 2000. The entire 16,569bp human mitochondrial genome is sequenced using established PCR primer pairs, which are known not to amplify pseudogenes, or other nuclear loci. Any putative DNA changes are confirmed by comparison to the revised "Cambridge" human mtDNA reference (Andrews et al. 1999). The sequences obtained from the tumour mtDNA are first compared for known polymorphisms (Andrews et al. 1999; MITOMAP) and then compared with the mtDNA sequence from the normal perilesional skin to identify genuine somatic mutations.

DHPLC is performed on the WAVE™ DNA Fragment Analysis System (Transgenomic, Omaha, USA) which provides a fully automated screening procedure. The same technology is used to screen for heteroplasmic mutations in the skin tumour mtDNA.

Using the back to back primer methodology described in Example 2, the pattern of DNA length mutations (i.e. tandem duplications) in the hypervariable segments of the non-coding region (NCR) are rapidly screened.

Example 4: deletion spectrum of the entire mitochondrial genome in human NMSC and its precursor lesions

MtDNA damage in squamous cell carcinomas (SCCS), Basal cell carcinomas (BCCS) and putative precursor lesions such as Bowen's disease and actinic keratoses (As) was compared to adjacent perilesional skin taken from different sun-exposed body sites. A long-extension PCR technique (LX-PCR) (Ray et al. 1998) was used to amplify the entire mitochondrial genome in order to determine the whole deletion spectrum of mtDNA. A myriad of specific deletions have been observed to occur in the mitochondrial genome. Not all deletions will correlate with non-melanoma skin cancer; however, for an accurate diagnostic method, those deletions that are associated with the disease must be known.

DNA is extracted by use of a commercial kit (Qiangen) according to the manufacturer's recommendations. The entire mitochondrial genome is amplified in two separate reactions using the Expand TM Long Template PCR System™ (Boehringer

Manheim, Switzerland). The PCR primers used are those described by Kleinle et al. (1997) covering the following regions of the Cambridge sequence (Andrews et al. 1999): DIA(nucleotides (nt) 336-363), DIB (nt 282-255), OLA (nt 5756-5781), and OLB (nt 5745-5781). These large products eliminate amplification of nuclear pseudogenes. The sequences of the primers are as follows:

DIAF: (336-363) 5' AACACATCTCTGCCAAACCCCAAAAACA 3'

OLBR: (5745-5721) 5' CCGGCGGCGGGAGAAGTAGATTGAA 3'

OLAF: (5756-5781) 5' GGGAGAAGCCCCGGCAGGTTTGAAGC 3'

DIBR: (282-255) 5' ATGATGTCTGTGTGGAAAGTGGCTGTGC 3'

Amplifications are performed in 50 microlitre reactions containing 16 pmol of each primer, 500μmol dNTPs, 10 x PCR buffer with 22.5mM $MgCl_2$ and detergents(kit), 0.75 μl of enzyme (3.5 x 10^3 units/ml) and 50-200ng of total DNA. One reaction generates 11,095bp segments of the genome, while another results in 5,409bp lengths (e.g. Kleinle et al, 1997). The PCR amplification conditions consists of a denaturing stage at 93°C for 1 min 30s, followed by 10 cycles of 93°C for 30s, 60°C for 30s and 68°C for 12 min, followed by a further 20 cycles of the same profile with an additional 5s added to the elongation time every cycle. There is a final cycle of 93°C for 30s, 60°C for 30s and an elongation time of 68°C for 26 minutes. To ensure reproducibility, a known amount of DNA is separated on a 1% agarose gel and only samples which have at least the same amount of DNA are included in the analysis.

A greater mean number of deletions is found with increasing UV exposure in the tumour samples, as shown in Table 2.

UV exposure	Mean number of deletions in adjacent normal epidermis	Mean number of deletions in epidermal tumour
Constant (n=5)	1.0	3.6
Intermittent (n=2)	0	1.5
Sun-protected (n=2)	0	0

Table 2. Comparison of the mean number of deletions observed in the LX-PCR of mtDNA between normal and tumour skin taken from different UV-exposed body sites.

Example 5: Aging and MtDNA

Using temporal maternal line comparisons (i.e. great-grandchild through great-grand parents), the entire sequence of mtDNA extracted from a given tissue is rapidly, and accurately sequenced, in order to definitively state the arrangement of nucleotide base pairs for that specific molecule and possible changes through time. These characterizations are compared to health status, aging indicators and between specific maternal lines, within larger populations. This combined information allows crucial statistical discrimination between separate causes resulting in the same mutation/deletion and establishes that the mtDNA sequences, used as a bio-marker, has the required index of specificity and sensitivity in order to establish its validity. In addition, the proportions of base pair deletions and mutations are compared for consistency in various tissues across the 4 maternal generations. Recent methodological developments have permitted detection of base pair deletions implicated in aging in blood samples (Bassam et al. 1991) and have raised the possibility that blood samples may be used to study mtDNA in lieu of skeletal muscle (von Wurmb et al. 1998). After establishing the efficacy of employing leukocytes in lieu of muscle tissue, as representative of mtDNA deletions and /or mutations, the next step measures only mtDNA in leukocytes. MtDNA deletions/mutations are then determined as previously described.

Skeletal muscle or leukocytes are obtained from a patient. DNA is extracted as set out in Example 1. The following primers were used:

12ST1: (1257-1279) 5' TATACCGCCATCTTCAGCAAAC3'

12ST2: (1433-1411) 5' TACTGCTAAATCCACCTTCGAC 3'

D1F: 5' CCTTACACTATTCCTCATCACC 3'

D1R: 5' TGTGGTCTTTGGAGTAGAAACC 3'

Amplifications were performed in 50 microlitre reactions containing 2.0 μmol of each primer, 250μmol dNTPs, 10 x PCR buffer(Thermopol Reaction Buffer), bovine

serum albumin, 0.5units Deep vent polymerase and 50-200ng of total DNA. The PCR amplification conditions consists of a denaturing stage at 95°C for 5 min (hot start), followed by 30 cycles of 94°C for 30s, 60°C for 60s and 72°C for 30s with a final extension at 72°C for 10 min. Gel electrophoresis was performed on a 2% agarose gel at 125 volts for 60 min, stained with ethidium bromide, and visualized under UV light. To ensure reproducibility, a known amount of DNA was separated on a 2% agarose gel and only samples which have the same amount of DNA were included in the analysis.

Example 6: Quantitative detection of the 4977bp common mtDNA deletion by 3-primer PCR

Where appropriate the incidence of the common deletion is determined in a quantitative manner by a 3-primer PCR method which detects levels greater than 1-5% or a dilution PCR method which detects levels less than 1 % down to 10^{-4}%. (See Example 7) Samples are obtained and DNA extracted as described in Example 1. To simultaneously detect and quantify the ratios of both deleted and wild type (wt) mtDNAs in the DNA samples, a 3-primer PCR procedure is used (as described in Birch-Machin et al 1998). Primers A, and C correspond to heavy strand positions 13720-13705 and 9028-9008 respectively (Anderson et al., 1981); primer B corresponds to light strand positions 8273-8289. Primer C maps to a mtDNA region within the common deletion, whereas primers A and B flank the deleted region. Therefore primers B and C only amplify wt-mtDNAs and primers A and B only amplify deleted mtDNAs (the distance between the two primers in the absence of the deletion, approximately 5.5kb, is too long to be amplified under our PCR conditions as described below).

Using three primers allowed the simultaneous detection of two bands, the larger one (755bp) corresponding to the wt-mtDNA, and the smaller one (470bp) corresponding to deleted mtDNA harbouring the 'common deletion'. The PCR reaction mixture (25µl total volume) contained 100ng total cellular DNA, 200µM dNTPs, 10mM Tris-HCl (pH 8.8), 50mM KCl, 1.5mM Mg Cl_2 , 0.1% Triton X-100, 2.5U *Taq* DNA polymerase (BioTaq, BiolineUK Limited, London), 25 pmoles of primers A and B, 6.25 pmoles of primer C and 3µCi of [α-^{32} P]-dATP. The PCR conditions were 25 cycles of 94°C at 1 minute, 55°C at 1 minute, 72°C at 2 minutes including a final extension of 15 minutes at 72°C. These PCR

products were then electrophoresed through a 6% nondenaturing polyacrylamide gel and the radioactive PCR fragments were quantified by phophorimage analysis using the ImageQuant™ software (Molecular Dynamics, Chesham UK).

Example 7: Serial Dilution PCR method to quantitatively detect low levels (<1%) of the common mtDNA deletion

A semi-quantitative PCR method (Corral-Debrinski *et al* 1991) is used to estimate the proportion of the common deletion in the total mtDNA extracted from the tissue/cell samples. Biological samples are obtained and DNA extracted as described in Example 1. The DNA sample is initially linearised using the restriction enzyme *Bam* HI (1µl enzyme and 1µl of commercially supplied buffer) at 37°C for 90 minutes. Serial dilutions are performed in two-fold steps (for total mtDNA there was an initial 10-fold dilution) and PCR performed for each dilution (1µl) using the following primers:

Primers for total mtDNA

L3108 (nt3108-3127)

H3717 (nt3717-3701)

Primers for Common Deletion

L8282 (nt8282-8305)

H13851(nt13851-13832)

The reaction conditions are as follows:

One cycle 94°C for 2 minutes, 34 cycles of 94°C for 45 seconds, 51°C for 30 seconds (total mtDNA), 56°C for 30 seconds (common deletion), 72°C for 1 minute and one final cycle of 72°C for 8 minutes. All PCR reactions are carried out in the following mixture (50µl): Sample DNA 1µl, 0.6µM forward primer, 0.6µM reverse primer, 0.2mM dNTP's, 5µl GeneAmp® 10x PCR Buffer, (Perkin Elmer), 0.2µl Amplitaq® DNA polymerase (Perkin Elmer), 35.75µl sterile autoclaved double distilled water.

Following electrophoresis the PCR productes are visualised on a UV transilluminator (TMW-20, Flowgen Ltd., Lichfield, UK) and a digital image of the gel obtained using image acquisition apparatus (Alpha Imager 2000, Alpha Innotech

Corporation, supplied by Flowgen Ltd., Lichfield, UK). The associated image analysis software (Alpha Ease v3.3, Alpha Innotech Corp.) allows the calculation of the integrated optical density (IOD) for each PCR product in a dilution series. The band where an IOD value of zero is obtained for both total mtDNA and deleted mtDNA and the corresponding dilution values are used to calculate the percentage of common deletion in the sample thus:

$$\%\text{common deletion} = \frac{\text{total mtDNA dilution factor}^{(\text{IOD Zero})} \times 100}{\text{common deletion dilution factor}^{(\text{IOD Zero})}}$$

Example 8: Denaturing high performance liquid chromatography (DHPLC)

Samples are obtained and DNA extracted as in Example 1. PCR in 13 overlapping fragments using two different PCR conditions as described by van den Bosch et al. (2000). The following three mtDNA specific primer pairs for PCR:

	i. Oligo Sequence
Mt3118F	CCCTGTACGAAAGGACAAGAG
Mt3334R	TGAGGAGTAGGAGGTTGG
Mt8207F	CCCATCGTCCTAGAATTAATTCC
Mt8400R	ATGGTGGGCCATACGGTAG
Mt14427F	CCCATGCCTCAGGATACTCCTC
Mt14997R	GCGTGAAGGTAGCGGATG

The 1-2 kb PCR products are digested into fragments of 90-600bp and resolved at their optimal melting temperature. Mutations are represented as two peaks and mutations with low percentages, such as <2% heteroplasmy as a "shoulder" in the peak.

DHPLC is performed with a mobile phase consisting of two eluents (pH 7.0). Buffer A contains triethylammonium acetate (TEAA), which interacts with both the negatively charged phosphate groups on the DNA as well as the surface of the column. Buffer B contains TEAA with 25% of the denaturing agent acetonitrile. Fragments were

eluted with a linear acetonitrile gradient at a constant flow rate. Increasing the concentration of acetonitrile will denature the fragments. Table 3 below is an example of a standard method for DHPLC of a PCR reaction generated using the WAVEMAKER software (Transgenomics) according to manufacturer's instructions.

5

Table 3: Standard Method for DHPLC

Step	Time	%A (buffer)	%B (buffer)	Ml/min (flow rate)
Loading	0.0	52	48	0.90
Start Gradient	0.1	47	53	
Stop Gradient	4.1	39	61	
Start Clean	4.2	0	100	
Stop Clean	4.7	0	100	
Start Equilibrate	4.8	52	48	
Stop Equilibrate	6.8	52	48	

The temperatures for successful resolution of the various heteroduplexes are detailed below and can simply be substituted into the relevant places in Table 2:

10

Fragment	Melting temp (°C)	Gradient of %Buffer B
Mt3118F	59	51-59
Mt8207F	58	50-58
Mt14427F	56	60-68

References

Anderson S, et al., Nature 290:457-464, 1981

Andrews RM, et al., *Nature Genetics* 23(2):147, 1999

Barringer et al., *Gene*, 89:117 1990

Bassam BJ, Caetano-Anolles PM, Gresshoff PM., Anal. Biochem. 196: 80-83, 1991

Berneburg M, et al., *J. Biol. Chem.* 274(22):15345-15349, 1999

Berthon P, Valeri A, Cohen-Akeninc A, Drelon E, Paiss T, Wohr G, Latil A et al., Am. J. Hum. Genet., 62: 1416-1424,1998

Birch-Machin MA, et al., Methods in Toxicology, Volume 2, 51-69,1993

Birch-Machin MA, et al., *J.Invest.Dermatol.*, 110:149-152 1998

Birch-Machin MA, Online Conference Report (Sunburnt DNA), International Congress of Biochemistry and Molecular Biology, New Scientist, 2000(a)

Birch-Machin MA, Taylor RW, Cochran B, Ackrell BAC, Tumbull DM. *Ann Neurol* 48: 330-335, 2000(b)

Birch-Machin MA, *Clin. Exp. Dermatol*,. 25(2), 141-146, 2000 (c)

Birch-Machin MA and Krishnan K. *Mitochondrion*, 1, p45 (2001).

Birch-Machin MA, Lindsey J. Lusher M and Krishnan K. *Mitochondrion*, 1 Suppl. 1, S30 (2001).

Bogliolo, M, et al., Mutagenesis, 14: 77-82, 1999

Brierley EJ, Johnson MA, Lightowlers RN, James O, Turnbull DM., Ann Neurol 43(2):217-223, 1998

Brockington, et al., *Nature Genet* 4:67-71, 1993

Brown, M.D., et al., Am J. Humn Genet, 60: 381-387, 1997

Brumley, R. L. Jr. and Smith, L.M., 1991, Rapid DNA sequencing by horizontal ultrathin gel electrophoresis, Nucleic Acids Res. 19:4121-4126 Nucleic Acids Res. 19: 4121-4126

Buttyan R, Sawczuk IS, Benson MC, Siegal JD, Olsson CA., Prostate 11:327-337,1987

Byrne E., Curr Opin Reumatol 4(6):784-793, 1992

Cairns P, Okami K, Halachmi S, Halachmi N, Esteller M, Herman JG, Jen J et al., Cancer Res 57:4997-5000, 1997

Chee, M. et al *Science* 274; 610-614, 1996

Chinnery PF, Howel N, Turnbull DM. *J.Med.Genet.*; 36: 425-436, 1999

Chinnery PF and Turnbull DM., Lancet 354 (supplement 1): 17-21, 1999

Chinnery PF and Turnbull DM., *Lancet* 354 (supplement 1): 17-21, 2000

Chollat-Traquet, C, Tobacco or health: a WHO programme., Eur J Cancer, 28(2-3): 311-315, 1992

Chomyn A, et al., Proc. Natl. Acad. Sci. USA 89(10):4221-4225, 1992

Cohen D, Barton G, The cost to society of smoking cessation., Thorax. 53(2): S38-42, 1998

Corral-Debrinski et al., *Mutat Res*, 275: 169-180, 1991

Cortopassi G. A. and Arnheim, H. Detection of a specific mitochondiral DNA deletion in tissues of older humans, Nucleic Acids Tes. 18, 6927-6933 1990

Cortopassi G, Wang E., Biochim Biophys Acta 1271(1):171-176,1995

Croteau DL, Stierum RH, Bohr VA, *Mutat Res* 434(3):137-148, 1999

Current Protocols in Molecular Biology

Davis RM, Boyd GM, Schoenborn CA, "Common courtesy" and the elimination of passive smoking. Results of the 1987 National Health Interview Survey. JAMA 263(16): 2208-10, 1990

Dong JT, Isaacs WB, Rinker-Schaeffer CW, Vukanovic J, Ichikawa T, Isaaca JT, Barrett JC., Science 268:884-886, 1995

Driezen P, Brown KS., Searchable database of questionnaire items from populations surveys of tobacco use in Canada: A summary report to the Ontario Tobacco Research Unit (Toronto, Ontario) 1999

Easton RD, Merriwether AD, Crews DE, and Ferrell RE., *Am. J. Hum. Genet.* 59:202-212, 1996

Fahn H, Wang L, Hseith R, Chang S, Kao S, Huang M, and Wei Y. *American Journal of Respiratory Critical Care Medicine*, 154:1141-1145, 1996

Fahn HJ, Wang LS, Kao SH, Chang SC, Huang MH, Wei YH., *Am. J. Respir. Cell. Mol. Biol.*, 19(6): 901-9, 1998

Finegold D., Mitochondrial Disease- Primary Care Physican's Guide. Psy-Ed. Corp D/B/A *Exceptional Parents Guide*: 12, 1997

Flanagan N, Birch-Machin MA, Rees JL., *Hum Mol Genet* 9 (17):2531-2537,2000

Flanagan N, Ray AJ, Todd C, Birch-Machin MA and Rees JL. J Invest. Dermatol (2001) 117 (5) 1314-1317

Fliss MS, et al. *Science* 287: 2017-2019, 2000

Gattermann, N, Berneburg, M, Heinisch, J, Aul, C, Schneider, W., *Leukemia* 9(10): 1704-10, 1995

Green R, Reed JC., *Science* 281 (5381):1309-1312, 1998

Guatelli, et al., *Proc. Nat. Acad. Sci.* U.S.A. 87: 1874 1990

Gulavita, Sunil Dr. Northwestern Ontario Cancer Centre – Personal Communication

Habano S, Nakamura, Sugai T., *Oncogene* 17 (15):1931-1937, 1998

Harding RM, et al., *Am. J. Hum. Genet.* 66, 1351-1361, 2000

Harman, D., *Proc Nati Acad Sci USA* 78(11): 7124-8, 1981

Hattori et al, Age-dependant increase in deleted mitochondrial DNA in the human heart: possible contributory factor to presbycardia, *AM. Heart J.*, 121, 1735-1742, 1991

Hayashi J, Ohta S, Kikuchi A, Takemitsu M, Goto Y, Nonaka 1., *Proc Natl Acad Sci USA* 88 (23):10614-10618, 1991

Hayashi J, Ohta S, Kikuchi A, Takemitsu M, Goto Y, Nonaka I., *Proc Natl Acad Sci USA* 88 (23):10614-10618, 1991

Hayward SW, Grossfeld GD, Tlsty TD, Cunha GR., *Int J Oncol* 13:35-47, 1998

Healy E, Birch-Machin MA, Rees JL. Chapter 1 1. *The Human Melanocortin 1 Receptor Gene. In the Melanocortin Receptors*(Cone RD (ed)). *Humana Press Inc.* New Jersey, USA, 1999

Healy E, Birch-Machin MA, Rees Jl., *Lancet* 355, 1072-1073, 2000

Hearst N, Hulley SB. Using secondary data, *In Designing clinical research: an epidemiological approach. Ed. Hulley, S. and Cummings, S., Baltimore: Williams & Wilkins,* pages 53-62, 1988

Hopgood, R., et al, 1992, Strategies for automated sequencing of human mtDNA directly from PCR products, *Biotechniques* 13:82-92

Hsieh, RH and Wei, YH, Age-dependent multiple deletions in human muscle mitochondrial DNA, in preparation 1992

Huang GM, Ng WL, Farkas J, He L, Liang HA, Gordon D, Hood R., *Genomics* 59(2):178-86,1999

Http://www.ornl.gov/hgmis/project/budget.html)

Ikebe et al., Increase of deleted mitochondrial DNA in the striatum in Parkinson's disease and senescence, *Biochem. Biophys. Res. Commun.* 170, 1044-1048, 1990

Innis et al PCR Protocols, *A Guide to Methods and Application*, Academic Press Inc. San Diego 1990

Kaiserman MJ, *Chronic Dis Can* 18(1): 13-9, 1997

Kalra J, Chaudhary AK, Prasad K,. *Int. J. Exp. Pathol.* 72(1): 1-7, 1991

Katayama et al., Deleted mitochondrial DNA in the skeletal muscle of aged individuals, *Biochem. Int.*, 25, 47-56 1991

Kleinle S, et al., *Human Genet.* 290: 457-465, 1997

Konishi N, Cho M, Yamamoto K, Hiasa Y. *Pathol. Int.* 47:735-747,1997

Krishnan K and Birch-Machin MA. British Journal of Dermatology (2002), 146,723

Kwoh et al. Proc. *Natl. Acad. Sci. U.S.A.*, 86: 1173 1989

Landis SH, Murray T, Bolden S, Wingo PA. *Cancer J. Clin.* 49:8-31

Landis, SH, et al., *CA Cancer J.. Clin.* 49: 8-31, 1999

Landegren et al. *Science*, 241: 1077 1988

LeDoux SP, et al. *Mutat Res* 434(3):149-159, 1999

Lee HC, et al. *FEBS Letters* 354:79-83,1994

Lee HC, Lu CY, Fahn HJ, Wei YHu. *Federation of European Biochemical Societies, 441:292*-296,1998

Lee HC, et al. *Arch. Biochem. Biophys.* 362(2): 309-16, 1999

Leonard & Shapira 1997

Lindsey J, Lusher M, Krishnan KJ and Birch-Machin MA., British Journal of Dermatology (2001), 144,655

Li Y, et al., In: *Oxygen Radicals and the Disease Process*, Amsterdam, The Netherlands: Harwood Academic Publishers, 237-277, 1997

Linnane et al., 1990

Liu CS, Kao SH, Wei YH. *Environ. Mol. Mutagen* 30(1): 47-55, 1997

Lowes S, Krishnan K, Lindsey J, Lusher M and Brich-Machin MA. British Journal of Dermatology (2002, 146,736

Luckey, J.A., et al, 1993, High speed DNA sequencing by capillary gel electrophoresis, *Methods Enzymol.* 218: 154-172

McCormack, Douglas. Website: http://cormactech.com/dna, 2001

Meibner C, von Wurmb N, Oehmichen M., *Int. J. Legal Med.* 110: 288-291, 1997

Michikawa Y, Mazzucchelli F, Bresolin N, Scarlato G, Attardi G., *Science* 286: 774-779,1999

Miquel J, de Juan E, Sevila I. EXS 62:47-57,1992

MITOMAP: A human mt genome database (www.gen.emory.edu/mitomap.html)

Mullis and Faloona *Methods Enzymol* 155, 335 1987

Nachman MW, Brown WM, Stoneking M, Aquardo CF., *Genetics* 142:53-963,1996

National Cancer Institute of Canada, Canadian cancer statistics 2000., National Cancer Institute of Canada., Toronto, Ont. 2000

Naviaux, RK., Mitochondrial Disease- Primary Care Physican's Guide. Psy-Ed. Corp D/B/A *Exceptional Parents Guide*: 3-10, 1997

Newton, CR and Graham, A., *Introduction toBiotecniques Series* 1997

Oefner PJ, Underhill PA., *Current protocols in human genetics* 19, 7.10.1-12, 1998

Ozen M, et al, *Prostate* 36:264-271,1998
Pang et al, *Arch. Biochem.Biophys.* 312: 534-538, 1994

Parsons TJ, et al., *Nature Genet.* 15 (4):363-368, 1997

Pascucci B, et al., *J.Mol.Biol.* 273 (2):417-427, 1997

Penta JS, Johnson FM, Wachsman JT, Copeland WC., Mut. Res. 488, 119-133, 2001

Polyak Y, et al., *Nature Genet.* 20 (3):291-293, 1998

Ray AJ, Rees JL, Birch-Machin MA., *J. Invest.Dermatol.* 110:692, 1998

Ray AJ, Rees JL, Birch-Machin MA., *Brit.J.Dermatol.*140:788, 1999

Ray AJ, Pickersgill L, Turner R, Nikaido O, Rees JL, Birch-Machin MA., *J. Invest. Dermatol* 115(4):674-679, 2000

Rees JL, Skin cancer. In: *The Genetic Basis of Human Cancer*, eds Vogelstein B, Kinzler K. New York: McGraw-Hill, pp527-536, 1998

Rehman I, Quinn AJ, Healy E, Rees JL. *Lancet* 344: 788-789, 1994

Rehman I, Takata M, Wu YY, Rees JL. *Oncogene* 12: 2483-2490, 1996

SAS Enterprise Mining Users Guide, SAS Inc., 2000

Sawyer E, Van Houten B., *Mutation Res:* 434(3):161-176, 1999

Schurr TG, Ballinger SW, Gan Y, Hodge JA, Merriwether DA, Lawrence DN, Knowler WC, Weiss KM, and Wallace DC., *Am. J. Hum. Genet.* 46:613-623, 1990

Seidman, M.D. et al., *Arch. Otolaryngol Head Neck Surg.*, 123: 1039-1045, 1997

Shankey TV, Jin JK, Dougherty S, Flanigan RC, Graham S, Pyle JM., *Cytometry* 21:30-39,1995

Shay JW, Werbin H., *Mutat. Res*:186: 149, 1987

Sherrat EJ, Thomas AW, Alcolado JC., *Clin. Sci.* 92:225-235,1997

Shoffner JM, Brown MD, Torroni A, Lott MT, Cabell MF, Mirra SS, Beal MF, Yang C, Gearing M, Salvo R, Watts RL, Juncos JL, Hansen LA, Crain BJ, Fayad M, Reckford CL, and Wallace DC., *Genomics* 17: 171-184, 1993

Smith DG, Malhi RS, Eshleman J, Lorenz JG and Kaestle FA., *Am. J. Hum. Genet.* 110:271-284, 1999

Smith R, Birch-Machin MA, Rees JL. *J. Invest. Dermatol.* 111: 101 -104, 1998
SpringNet – CE Connection: Screening, Diagnosis: Improving Primary Care Outcomes. Website: http://www.springnet.com/ce/j803a.htm

Stone AC and Stoneking M. Amer. J., *Phys. Anthro.* 92(4):463-471, 1993

Tamura S, et al. *Eur.J.Cancer [A]* 35 (2):316-319, 1999

Tanaka, M. et al, 1996, Automated sequencing of mtDNA, *Methods Enzymol.* 264: 407-421

Taniike, M. et al., *BioChem BioPhys Res Comun*, 186: 47-53, 1992

Taylor RW, Birch-Machin MA, Bartlett K, Turnbull DM., *J Biol Chem*, 269, 3523-3528 1994

Tijssen, P. (ed) *Laboratory Techniques in Biochemistry and Molecular Biology,* Vol. 24: *Hybridization with Polynucleotide Probes,* Elsevier, N.Y., 1993

Tori et al., Ageing-associated deletions of human diaphragmatic mitochondrial DNA, *AM. J. Respir. Cel.l Mol. Biol.* in press 1992

Van De Graff, KM, Fox, SI. *Concepts of Human Anatomy and Physiology.*, Dubuque: WM. C. Brown Publishers, 1995

Van den Bosch BJC, et al., *Nucleic Acids Res.* 28: 89, 2000

von Wurmb, N, Oehmichen, M, Meissner, C., *Mutat Res.* 422:247-254, 1998

Wallace DC., *Annu Rev Biochem,* 61: 1175-1212,1992

Wallace DC. *Proc. Natl. Acad. Sci. USA 91: 8739-8746, 1994*

Wald and Wallace, D.C., Mitochondrial Diseases in man and Mouse. *Science*, 5(283): 1482-1497, 1999

Wald, NJ, Hackshaw,AK, Cigarette Smoking: an epidemiological overview. *Br Med Bull.* 52(1): 3-11, 1996

Wallace et al., Mitochondiral DNA MUtatio Assoicated with Leber's Hereditary Optic Neuropathy, *Science,* 1427-1429

Walsh PC, Partin AW. *Cancer* 80:1871-1874,1997

Ward RH, Frazier BL, Dew-Jager K, Paabo S., *Proc. Natl. Acad. Sci.* USA 88:8720-8724, 1991
Ward 1993

Wei YH, Pang C, You B, Lee H. *Ann.NY Acad.Sci.* 786:82-101, 1996

Wei YH. Proceedings of the Nat. Sci. Council of the Republic of China April 22(2):5567, 1998

Weinstock MA: In: JJ Stern RS, MacKie RM and Weinstock MA, Grob *(eds) Epidemiology,* Blackwell (UK). pp121-128, 1998

Woodwell DA. National Ambulatory Medical Care Survey: 1997 Summary. Advance data from vital and health statistics; no. 305. Hyattsville, Maryland: National Center for Health Statistics. 1999

Wu & Wallace *Genomics,* 4:560, 1989

Xu J, et al., *Nature Genet* 20: 175-179,1998

Yamaguchi KT, et al., *Free Radical Res. Commun.* 16(3):167-74, 1992

Yeh, J.J., et al., *Oncogene Journal*, 19: 2060-2066, 2000

Yen et al., Ageing-associated 5kb deletion in human liver mitochondrial DNA, *Biochem., Biophys., Res. Commun.*, 178, 124-131 1991

Yen et al., Age-dependent 6 kb deletion in human liver mitochondirial DNA, *Biochem. Int.* 26, 457-468 1992

Zeviani M, et al. *Am. J. Hum. Genet.* 47:904-914,1990

Zhang et al., Multiple mitochondiral DNA deletions in an elderly human individual, *FEBS Lett*, 297, 34-38 1992

Zhang, C., et al., *BioChem. BioPhys. Res. Comun.*, 195: 1104-1110, 1993

WE CLAIM:

1. A method of detecting in a subject containing mtDNA the genesis or progression of disease comprising:

a) obtaining a biological sample from the subject;

b) extracting DNA from the biological sample;

c) detecting the presence of mutations in the mtDNA; and

d) comparing the mtDNA of the biological sample to a database, the database containing data of mutations associated with the mitochondrial DNA sequences of non-disease and disease associated mitochondrial genomes.

2. The method of claim 1 wherein the step of detecting the presence of mutations is selected from the group consisting of:

a) sequencing the mtDNA;

b) amplifying mtDNA by PCR;

c) Southern, Northern, Western and South-Western blot hybridizations;

d) denaturing HPLC;

e) hybridization to microarrays, gene chips or biochips;

f) molecular marker analysis; and

g) a combination of any of a) through f).

3. The method of claim 2 where the mitochondrial DNA which is sequenced comprises specific areas of the mitochondrial genome where known biomarkers associated with disease are located.

4. The method of claim 2 where the mitochondrial DNA which is sequenced comprises the entire mitochondrial genome.

5. The method of claim 1 where the biological sample is from a tissue suspected of being a potential site of disease.

6. The method of claim 1, where the biological sample is from a tissue suspected of harbouring a metastasis.

7. The method of claim 1 where the disease is selected from the group of prostate cancer or non-melanoma skin cancer.

8. The method of claim 7, where the disease is prostate cancer.

9. The method of claim 7, where the disease is non-melanoma skin cancer.

10. The method of claim 1, where the mutations is/are selected from the group of single base pair mutations, deletions, insertions, and transversions.

11. The method of claim 1 where the mutation is homoplasmic.

12. The method of claim 1 where the mutation is heteroplasmic at any level.

13. The method of claim 1, where the biological sample is from the group selected from blood, sputum, buccal cells, saliva, prostate massage fluid, sweat, cervical tissue from a PAP smear, urine, skin cells, bone, hair, lymph tissue, cervical smears, breast aspirate, fecal matter, ejaculate, menstrual flow or biopsy tissue.

14. The method of any claims 1 to 13, where the database contains at least a statistically significant number of mitochondrial DNA sequences, the mitochondrial DNA sequences having been obtained from both maternal line and non-maternal line samples.

15. A method of detecting in a subject containing mtDNA the presence of a disease comprising:

a) obtaining a biological sample from the subject;
b) extracting DNA from the biological sample;
c) detecting the presence of mutations in the mtDNA; and
d) comparing the mtDNA of the biological sample to a database, the database containing data of mutations associated with the mitochondrial DNA sequences of non-disease and disease associated mitochondrial genomes.

16. The method of claim 15 wherein the step of detecting the presence of mutations is selected from the group consisting of:

a) sequencing the mtDNA ;
b) amplifying mtDNA by PCR;
c) Southern, Northern, Western and South-Western blot hybridizations;
d) denaturing HPLC;
e) hybridization to microarrays, gene chips or biochips;
f) molecular marker analysis; and
g) a combination of any of a) through f).

17. The method of claim 16 where the mitochondrial DNA which is sequenced comprises specific areas of the mitochondrial genome where known biomarkers associated with disease are located.

18. The method of claim 16 where the mitochondrial DNA which is sequenced comprises the entire mitochondrial genome.

19. The method of claim 15 where the biological sample is from a tissue suspected of being a potential site of disease.

20. The method of claim 15, where the biological sample is from a tissue suspected of harbouring a metastasis.

21. The method of claim 15 where the disease is selected from the group of prostate cancer or non-melanoma skin cancer.

22. The method of claim 21, where the disease is prostate cancer.

23. The method of claim 21, where the disease is non-melanoma skin cancer.

24. The method of claim 15, where the mutations is/are selected from the group of single base pair mutations, deletions, insertions, and transversions.

25. The method of claim 15 here the mutation is homoplasmic.

26. The method of claim 15 where the mutation is heteroplasmic at any level.

27. The method of claim 15, where the biological sample is from the group selected from blood, sputum, buccal cells, saliva, prostate massage fluid, sweat, cervical tissue from a PAP smear, urine, skin cells, bone, hair, lymph tissue, cervical smears, breast aspirate, fecal matter, ejaculate, menstrual flow or biopsy tissue.

28. The method of any claims 15 to 27, where the database contains at least a statistically significant number of mitochondrial DNA sequences, the mitochondrial DNA sequences having been obtained from both maternal line and non-maternal line samples.

29. A method of determining a predisposition to a disease or disorder indicated by mutations in a mitochondrial DNA sequence comprising:
a) obtaining a biological sample from the human subject;
b) extracting DNA from the biological sample;
c)detecting the presence of mutations in the mtDNA; and
d) comparing the mtDNA of the biological sample to a database, the database containing data of mutations associated with the mitochondrial DNA sequences of non-disease and disease associated mitochondrial genomes.

30. The method of claim 29 wherein the step of detecting the presence of mutations is selected from the group consisting of:

a) sequencing the mtDNA ;
b) amplifying mtDNA by PCR;
c) Southern, Northern, Western and South-Western blot hybridizations;
d) denaturing HPLC;
e) hybridization to microarrays, gene chips or biochips;

f) molecular marker analysis; and

g) a combination of any of a) through f).

31. The method of claim 30 where the mitochondrial DNA which is sequenced comprises specific areas of the mitochondrial genome where known biomarkers associated with disease are located.

32. The method of claim 30 where the mitochondrial DNA which is sequenced comprises the entire mitochondrial genome.

33. The method of claim 29 where the biological sample is from a tissue suspected of being a potential site of disease.

34. The method of claim 29, where the biological sample is from a tissue suspected of harboring a metastasis.

35. The method of claim 29 where the disease is selected from the group of prostate cancer or non-melanoma skin cancer.

36. The method of claim 35, where the disease is prostate cancer.

37. The method of claim 35, where the disease is non-melanoma skin cancer.

38. The method of claim 29, where the mutation is selected from the group of single base pair mutations, deletions, insertions, and transversions.

39. The method of claim 29, where the mutation is homoplasmic.

40. The method of claim 29, where the mutation is heteroplasmic at any level.

41. The method of claim 29, where the biological sample is from the group selected from blood, sputum, buccal cells, saliva, prostate massage fluid, sweat, breast aspirate, fecal

matter, ejaculate, menstrual flow, cervical tissue from a Pap smear, urine, skin cells, bone, hair, lymph tissue, cervical smears or biopsy tissue.

42. The method of any of claims 29 to 41 where the database contains at least a statistically significant number of mitochondrial DNA sequences, the mitochondrial DNA sequences having been obtained from both maternal line and non-maternal line samples.

43. A method for assessing the status of the aging process of a human subject comprising:

a) obtaining a biological sample from the human subject;

b) extracting DNA from the biological sample;

c) detecting mutations in the mtDNA; and

d) comparing the mtDNA of the biological sample to a database, the database containing data of mutations associated with the mitochondrial DNA sequences of non-disease and disease associated mitochondrial genomes.

44. The method of claim 43 wherein the step of detecting the presence of mutations is selected from the group consisting of:

a) sequencing the mtDNA ;

b) amplifying mtDNA by PCR;

c) Southern, Northern, Western and South-Western blot hybridizations;

d) denaturing HPLC;

e) hybridization to microarrays, gene chips or biochips;

f) molecular marker analysis; and

g) a combination of any of a) through f).

45. The method of claim 43 where the mitochondrial DNA which is sequenced comprises specific areas of the mitochondrial genome where known biomarkers associated with disease are located.

46. The method of claim 43 where the mitochondrial DNA which is sequenced comprises the entire mitochondrial genome.

47. The method of claim 43, where the mutation is selected from the group of single base pair mutations, deletions, insertions, and transversions.

48. The method of claim 43 where the mutation is homoplasmic.

49. The method of claim 43 where the mutation is heteroplasmic at any level.

50. The method of any of claim 43, where the biological sample is from the group selected from blood, sputum, buccal cells, saliva, prostate massage fluid, sweat cervical tissue from a PAP smear, urine, skin cells, bone, hair, lymph tissue, cervical smears, fecal matter, breast aspirate, ejaculate, menstrual flow or biopsy tissue.

51. The method of any of claims 43 to 50 where the database contains at least a statistically significant number of mitochondrial DNA sequences, the mitochondrial DNA sequences having been obtained from both maternal line and non-maternal line samples.

52. A database containing a plurality of human mitochondrial DNA sequences, the mitochondrial DNA sequences selected from the group of normal control sequences associated with non-disease states, sequences associated with the presence of disease or sequences indicative of the predisposition to disease.

53. The database of claim 52, wherein the disease is selected from the group of prostate cancer and non-melanoma skin cancer.

54. The database of claim 52 where the mitochondrial DNA sequences are associated with the aging process of a human subject.

55. The database of claims 53 or 54 wherein the database contains a statistically significant number of mitochondrial DNA sequences, the mitochondrial DNA sequences having been obtained from both maternal line and non-maternal line samples.

56. A kit for diagnosis of a disease comprising a disposable chip, microarray, means for holding the disposable chip, means for extraction of mitochondrial DNA and means for access to a database of mitochondrial DNA sequences.

57. A kit for determining predisposition to a disease comprising a disposable chip, microarray, means for holding the disposable chip, means for extraction of mitochondrial DNA and means for access to a database of mitochondrial DNA sequences.

58. An array comprising a plurality of nucleic acid members, and a solid substrate, wherein each nucleic acid member is indicative of the presence of a disease and is selected from the group of mitochondrial DNA, RNA transcribed from mitochondrial DNA , wherein each nucleic acid member has a unique position on said array and is stably associated with the solid substrate.

59. The array of claim 58, wherein each member is indicative of prostate cancer.

60. The array of claim 58, wherein each member is indicative of non-melanoma skin cancer.

61. An array comprising a plurality of nucleic acid members, and a solid substrate, wherein each nucleic acid member is indicative of the predisposition to a disease and is selected from the group of mitochondrial DNA, RNA transcribed from mitochondrial DNA, and cDNA wherein each nucleic acid member has a unique position on said array and is stably associated with the solid substrate.

62. The array of claim 58, wherein each member is indicative of a predisposition to a disease.

63. The array of claim 58, wherein each member is associated with the aging process.

64. A method of diagnosing a disease in a patient comprising hybridizing a nucleic acid sample obtained from mitochondrial DNA to an array comprising a solid substrate and a

plurality of nucleic acid members, wherein each member is indicative of the presence of a disease, wherein each nucleic acid member has a unique position and is stably associated with the solid substrate, and wherein hybridization of said nucleic acid sample to one or more nucleic acid members comprising said array is indicative of the presence of the disease.

65. The method of claim 64, wherein the disease is prostate cancer.

66. The method of claim 65, further comprising the step of isolating a prostate massage fluid sample from said patient.

67. The method of claim 65, further comprising the step of preparing a nucleic acid sample from said prostate massage fluid sample.

68. A method of diagnosing non-melanoma skin cancer in a patient comprising: hybridizing a nucleic acid sample obtained from mitochondrial DNA to an array comprising a solid substrate and a plurality of nucleic acid members, wherein each member is indicative of non-melanoma cancer, wherein each nucleic acid member has a unique position and is stably associated with the solid substrate, and wherein hybridization of said nucleic acid sample to one or more nucleic acid members comprising said array is indicative of the presence of non-melanoma skin cancer.

69. The method of claim 68, further comprising the step of isolating a skin sample from said patient.

70. The method of claim 69, further comprising the step of preparing a nucleic acid sample from said skin sample.

71. A method of detecting heteroplasmy in a subject containing mtDNA comprising:

a) obtaining a biological sample from the subject;
b) extracting DNA from the biological sample; and
c) performing denaturing HPLC on the sample.

72. A method of detecting mutations associated with disease in a subject containing mtDNA comprising:

a) obtaining a biological sample from the subject;

b) extracting DNA from the biological sample;

c) detecting the presence of mutations in the mtDNA; and

d) comparing the mt DNA of the biological sample to a database, the database containing data of common population variants in non-disease and disease associated mitochondrial genomes.

专利名称(译)	完成线粒体基因组序列作为健康科学的诊断工具		
公开(公告)号	EP1397508A2	公开(公告)日	2004-03-17
申请号	EP2002737683	申请日	2002-06-10
[标]申请(专利权)人(译)	1304854 ONTARIO		
申请(专利权)人(译)	1304854 ONTARIO LTD		
当前申请(专利权)人(译)	1304854 ONTARIO LTD		
[标]发明人	BIRCH MACHIN MARK DAKUBO GABRIEL D PARR RYAN THAYER ROBERT NGOM ALIOUNE		
发明人	BIRCH-MACHIN, MARK DAKUBO, GABRIEL, D. PARR, RYAN THAYER, ROBERT NGOM, ALIOUNE TH'NG, JOHN,NW ONTARIO REGIONAL CANCER CENTRE		
IPC分类号	G01N33/53 C12M1/00 C12N15/09 C12Q1/68 G01N30/88 G01N33/566 G01N37/00		
CPC分类号	C12Q1/6886 C12Q2600/156		
优先权	60/297340 2001-06-11 US 2450004 2003-12-18 CA		
外部链接	Espacenet		

摘要(译)

对整个线粒体基因组中的突变的检查用作诸如前列腺癌和非黑素瘤皮肤癌的疾病的诊断系统。在线粒体基因组中的特征性突变和重排（包括点突变（转换，颠换），缺失，倒位，重复，重组，插入或其组合）用作前列腺癌和非黑素瘤皮肤癌的早期指标。此外，4977bp，或“共同缺失”以及其他相关的突变和/或缺失被用作衰老的量度。